

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 06:18:21 ; Search time 28.74 Seconds
(Without alignments)
1147.773 Million cell updates/sec

Title: US-09-237-981-10

Perfect score: 480

Sequence: 1 MKHLVAAMLLVGLSLGVPQF.....WSWYGRITLRSLLGCAEE 480

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 198801 seqs, 68722935 residues

Word size : 0

Total number of hits satisfying chosen parameters: 198801

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 50 summaries

Database : PIR.67:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	10	2.1	409	2 T11743	pp47 protein - pig
2	10	2.1	427	2 JC4915	ags protein precursor - rat
3	10	2.1	1220	2 A56136	jagged protein precursor
4	10	2.1	2555	2 A40043	notch protein homolog
5	9	1.9	218	2 A47285	milk fat globule protein
6	9	1.9	401	2 S65138	glycoprotein antigen
7	9	1.9	427	2 S74211	PAS-6/7 protein precursor
8	9	1.9	463	1 A36479	milk fat globule protein
9	9	1.9	1408	2 S16148	gene serrate protein
10	9	1.9	2139	2 A35672	crumbs protein - f
11	9	1.9	2319	2 A47004	coagulation factor
12	9	1.9	2437	2 S42612	transmembrane protein
13	9	1.9	2531	2 S18188	notch protein homolog
14	9	1.9	2531	2 A46019	Notch-1 protein - human
15	8	1.7	63	2 F75371	hypothetical protein
16	8	1.7	247	2 E75609	amino acid ABC transporter
17	8	1.7	330	2 T46256	brevican - human
18	8	1.7	481	2 T48516	probable oligopeptide
19	8	1.7	832	2 A31246	neurogenic protein
20	8	1.7	833	2 S19087	gene delta protein
21	8	1.7	880	2 S00670	neurogenic repeat
22	8	1.7	883	2 S49126	brevican precursor
23	8	1.7	883	2 S57653	brevican precursor
24	8	1.7	912	2 A54423	brevican precursor
25	8	1.7	1193	2 T21133	hypothetical protein
26	8	1.7	1251	2 A57293	latent transformin
27	8	1.7	1438	2 A48216	neurexin III-alpha
28	8	1.7	1471	2 B48218	neurexin III-alpha
29	8	1.7	1578	2 I48216	neurexin III-alpha

30	8	1.7	1715	2 C40228	neurexin II-alpha
31	8	1.7	2183	2 T42764	coagulation factor
32	8	1.7	2524	2 A35844	notch protein - Af
33	8	1.7	2531	2 T31070	notch homolog - se
34	8	1.7	2670	2 T37919	GCN1 homolog - fis
35	7	1.5	15	2 D48394	major fat-globule
36	7	1.5	96	2 T26872	hypothetical protein
37	7	1.5	106	2 A29760	hypothetical protein
38	7	1.5	107	2 C81232	50S ribosomal protein
39	7	1.5	126	2 B83265	hypothetical protein
40	7	1.5	135	2 C70890	hypothetical protein
41	7	1.5	149	1 CYFGA2	alpha-crystallin c
42	7	1.5	167	1 CYFGAA	alpha-crystallin c
43	7	1.5	171	2 F83332	hypothetical protein
44	7	1.5	172	2 B71464	hypothetical protein
45	7	1.5	172	2 C81726	conserved hypothetical
46	7	1.5	173	1 CYHUA	alpha-crystallin c
47	7	1.5	173	1 CYCHAA	alpha-crystallin c
48	7	1.5	173	1 CYEHAA	alpha-crystallin c
49	7	1.5	173	1 CYAQAA	alpha-crystallin c
50	7	1.5	173	1 CYLZAA	alpha-crystallin c

ALIGNMENTS

RESULT 1

T11743
pp47 protein - pig (fragment)
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C:Accession: T11743
R:Ensslin, M.; Vogel, T.; Calvete, J.J.; Thole, H.H.; Schmidtke, J.; Matsuda, T.; To; Biol. Reprod. 58, 1057-1064, 1998
A:Title: Molecular cloning and characterization of P47, a novel boar sperm-associated
A:Reference number: Z17325; MUID:98206817
A:Accession: T11743
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-409 <ENS>
A:Cross-references: EMBL:Y11683; NID:g2652927; PIDN:CAA72379.1; PID:g2652928
A:Experimental source: testis
C:Function:
A:Description: may be involved in membrane remodeling and/or function as a zona pellu
C:Superfamily: unassigned EGF-related proteins; EGF homology
F:6-40/Domain: EGF homology <EGF>

Query Match 2.1%; Score 10; DB 2; Length 409;
Best Local Similarity 100.0%; Pred. No. 0.031;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 310 ELLGCELSGC 319

|||||||

Db 243 ELLGCELSGC 252

RESULT 2

JC4915
ags protein precursor - rat
N:Alternate names: O-acetyl-Gd3 ganglioside
C:Species: Rattus norvegicus (Norway rat)
C:Date: 26-Sep-1996 #sequence_revision 01-Nov-1996 #text_change 20-Jun-2000
C:Accession: JC4915
R:Ogura, K.; Nara, K.; Watanabe, Y.; Kohno, K.; Tai, T.; Sanai, Y. Biochem. Biophys. Res. Commun. 225, 932-938, 1996
A:Title: Cloning and expression of cDNA for O-acetylation of GD3 ganglioside.
A:Reference number: JC4915; MUID:96374422
A:Accession: JC4915
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-427 <OGU>
A:Cross-references: DBJ:D84068; NID:g1620006; PIDN:BAA12210.1; PID:g1620007

F;1927-1959/Domain: ankyrin repeat homology <AN1>
F;1960-1992/Domain: ankyrin repeat homology <AN2>
F;1994-2026/Domain: ankyrin repeat homology <AN3>
F;2027-2059/Domain: ankyrin repeat homology <AN4>
F;2060-2092/Domain: ankyrin repeat homology <AN5>

Query Match 2.1%; Score 10; DB 2; Length 2555;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 NPCNGGTGCE 90
|||||
DB 684 NPCNGGTGCE 693

RESULT 5
A47285
milk fat globule protein - human (fragment)
C:Species: Homo sapiens (man)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 28-May-1999
R;Accession: A47285
R;Larocca, D.; Peterson, J.A.; Urtea, R.; Kuniyoshi, J.; Bistrain, A.M.; Ceriani, R.L.
Cancer Res. 51, 4994-4998, 1991
A:title: A Mr 46,000 human milk fat globule protein that is highly expressed in human
A;Reference number: A47285; MUID:91371351
A;Accession: A47285
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-218 <LAR>
C;Cross-references: GB:S56151; NID:g235396; PIDN:AAB19771.1; PID:g235397
A;Superfamily: milk fat globule protein; discoidin I amino-terminal homology; EGF hom
F;1-56/Domain: discoidin I amino-terminal homology (fragment) <DN1>
F;60-218/Domain: discoidin I amino-terminal homology <DN2>

Query Match 1.9%; Score 9; DB 2; Length 218;
Best Local Similarity 100.0%; Pred. No. 0.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 388 VTGIIQGA 396
|||||
DB 130 VTGIIQGA 138

RESULT 6
S65138
glycoprotein antigen MGP57/53, mammary gland - bovine (fragment)
N;Alternate names: glycoprotein component 16/major fat-globule membrane protein/MFG-E
C:Species: Bos primigenius taurus (cattle)
C;Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 07-Aug-1998
C;Accession: S65138; G48394
R;Aoki, N.; Kishi, M.; Taniguchi, Y.; Adachi, T.; Nakamura, R.; Matsuda, T.
Biochim. Biophys. Acta 1245, 385-391, 1995
A:title: Molecular cloning of glycoprotein antigens MGP57/53 recognized by monoclonal
A;Reference number: S65138; MUID:96125736
A;Accession: S65138
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-401 <AOK>
R;Mather, I.H.; Banghart, L.R.; Lane, W.S.
Biochem. Mol. Biol. Int. 29, 545-554, 1993
A:title: The major fat-globule membrane proteins, bovine components 15/16 and guinea-
II-like sequences.
A;Reference number: A48394; MUID:93250576
A;Accession: G48394
A;Status: preliminary
A;Molecule type: protein
A;Residues: 207-220 <MAP>
A;Experimental source: milk
A;Note: sequence extracted from NCBI backbone (NCBIP:131457)
C;Superfamily: milk fat globule protein; discoidin I amino-terminal homology; EGF hom
C;Keywords: glycoprotein

F:1-32/Domain: EGF homology (fragment) <EG1>
F:40-79/Domain: EGF homology <EG2>
F:82-239/Domain: discoidin I amino-terminal homology <DN1>
F:243-401/Domain: discoidin I amino-terminal homology <DN2>

Query Match 1.9%; Score 9; DB 2; Length 401;
Best Local Similarity 100.0%; Pred. No. 0.35;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 388 VTGIITOGA 396
|||||
DB 313 VTGIITOGA 321

RESULT 7

PAS-6/7 protein precursor - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 04-Dec-1997 #sequence_revision 12-Dec-1997 #text_change 20-Jun-2000
C:Accession: S74211; S78114; S24181
R:Hvarregaard, J.; Andersen, M.H.; Berglund, L.; Rasmussen, J.T.; Petersen, T.E.
Eur. J. Biochem. 240, 628-636, 1996
A:Title: Characterization of glycoprotein PAS-6/7 from membranes of bovine milk fat globule precursor
A:Reference number: S74211; MUID:97008954
A:Accession: S74211
A:Molecule type: mRNA
A:Residues: 1-427 <HVA>
A:Cross-references: EMBL:X91895; NID:gl632778; PIDN:CAA62997.1; PID:gl632779
A:Accession: S78114
A:Molecule type: protein
A:Residues: 19-85;96-110;140-165;174-216;221-232;248-277;285-293;309-337;339-420;425-427
R:Kim, D.H.; Kanno, C.; Mizokami, Y.
Biochim. Biophys. Acta 1122, 203-211, 1992
A:Title: Purification and characterization of major glycoproteins, PAS-6 and PAS-7, from membranes of bovine milk fat globule precursor
A:Reference number: S23926; MUID:92353107
A:Accession: S24181
A:Molecule type: protein
A:Residues: 383-394 <KIM>
C:Superfamily: milk fat globule protein; discoidin I amino-terminal homology; EGF homology
C:Keywords: blocked amino end; disulfide bond; glycoprotein; milk
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-427/Product: PAS-6/7 protein #status experimental <MAT>
F:24-58/Domain: EGF homology <EG1>
F:66-105/Domain: EGF homology <EG2>
F:108-265/Domain: discoidin I amino-terminal homology <DN1>
F:269-427/Domain: discoidin I amino-terminal homology <DN2>
F:24-35, 29-47, 49-58, 66-77, 71-94, 96-105/Disulfide bonds: #status predicted
F:27/Binding site: carbohydrate (Ser) (covalent) #status experimental
F:34/Binding site: carbohydrate (Thr) (covalent) #status experimental
F:59,227/Binding site: carbohydrate (Asn) (covalent) #status experimental
F:109-265,252-256,270-427/Disulfide bonds: #status experimental

Query Match 1.9%; Score 9; DB 2; Length 427;
Best Local Similarity 100.0%; Pred. No. 0.37;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 388 VTGIITOGA 396
|||||
DB 339 VTGIITOGA 347

RESULT 8

milk fat globule membrane protein - mouse
C:Species: Mus musculus (house mouse)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A36479
R:Stubbs, J.D.; Lekutis, C.; Singer, K.L.; Bui, A.; Yuzuki, D.; Srinivasan, U.; Parry, C.
Proc. Natl. Acad. Sci. U.S.A. 87, 8417-8421, 1990
A:Title: cDNA cloning of a mouse mammary epithelial cell surface protein reveals the existence of a novel protein family
A:Reference number: A36479; MUID:91046008

A:Accession: A36479
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-463 <STU>

C:Cross-references: GB:M38337; NID:gl99142; PIDN:AAA39534.1; PID:gl99143
C:Superfamily: milk fat globule protein; discoidin I amino-terminal homology; EGF homology
C:Keywords: membrane protein
F:28-60/Domain: EGF homology <EG1>
F:68-107/Domain: EGF homology <EG2>
F:147-303/Domain: discoidin I amino-terminal homology <DN1>
F:307-463/Domain: discoidin I amino-terminal homology <DN2>

Query Match 1.9%; Score 9; DB 1; Length 463;
Best Local Similarity 100.0%; Pred. No. 0.4;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 388 VTGIITOGA 396
|||||
DB 377 VTGIITOGA 385

RESULT 9

S16148
C:Species: Drosophila melanogaster
C:Date: 31-Dec-1991 #sequence_revision 02-Aug-1994 #text_change 17-Nov-2000
C:Accession: S16148; S16878; A36666
R:Thomas, U.; Speicher, S.A.; Knust, E.
Development 111, 749-761, 1991
A:Title: The Drosophila gene Serrate encodes an EGF-like transmembrane protein with a signal sequence
A:Reference number: S16148; MUID:91347903
A:Accession: S16148
A:Molecule type: mRNA
A:Residues: 1-1408 <THO1>
A:Cross-references: EMBL:X56811
R:Thomas, U.
submitted to the EMBL Data Library, November 1990
A:Reference number: S16878
A:Accession: S16878
A:Molecule type: mRNA
A:Residues: 1-1351, 'T', 1353-1408 <THO2>
A:Cross-references: EMBL:X56811; NID:g8563; PID:g8564
R:Fleming, R.J.; Scottgale, T.N.; Diederich, R.J.; Artavanis-Tsakonas, S.
Genes Dev. 4, 2188-2201, 1990
A:Title: The gene Serrate encodes a putative EGF-like transmembrane protein essential for morphogenesis of the wing
A:Reference number: A36666; MUID:91099666
A:Accession: A36666
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-15, 20-26, 'A', 28-1408 <FLE>
A:Cross-references: GB:M35759; NID:gl58605; PID:gl58606
C:Genetics:

A:Gene: FlyBase:Ser
A:Cross-references: FlyBase:FBgn0004197
C:Superfamily: unassigned EGF-related proteins; EGF homology
C:Keywords: glycoprotein; transmembrane protein
F:1-84/Domain: signal sequence #status predicted <SIG>
F:85-1408/Product: gene serrate protein #status predicted <MAT>
F:85-1221/Domain: extracellular #status predicted <EXT>
F:283-316/Domain: EGF homology <EG01>
F:319-348/Domain: EGF homology <EG02>
F:355-388/Domain: EGF homology <EG03>
F:395-488/Domain: EGF homology #status atypical <EG04>
F:493-526/Domain: EGF homology <EG05>
F:533-608/Domain: EGF homology #status atypical <EG06>
F:615-645/Domain: EGF homology <EG07>
F:652-683/Domain: EGF homology <EG08>
F:690-720/Domain: EGF homology <EG09>
F:727-796/Domain: EGF homology #status atypical <EG10>
F:803-834/Domain: EGF homology <EG11>
F:841-876/Domain: EGF homology <EG12>
F:883-914/Domain: EGF homology <EG13>

F;921-952/Domain: EGF homology <EG14>
 F;997-1060/Region: cysteine-rich
 F;1222-1246/Domain: transmembrane #status predicted <TM1>
 F;1247-1408/Domain: intracellular #status predicted <INT>
 F;152,196,247,331,412,452,558,739,965,977,1004,1030,1150/Binding site: carbohydrate (Asn)

Query Match 1.9%; Score 9; DB 2; Length 1408;
 Best Local Similarity 100.0%; Pred. No. 1.1;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 82 PCHNGTCE 90
 |||||
 Db 887 PCHNGTCE 895

RESULT 10
 A35672
 crumbs protein - fruit fly (Drosophila melanogaster)
 C:Species: Drosophila melanogaster
 C:Date: 21-Sep-1990 #sequence_revision 18-Nov-1992 #text_change 11-Jan-2000
 C:Accession: A35672
 R:Tepas, U.; Theres, C.; Knust, E.
 Cell 61, 787-799, 1990
 A:Title: crumbs encodes an EGF-like protein expressed on apical membranes of Drosophila
 A:Reference number: A35672; MUID:90263104
 A:Accession: A35672
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-2139 <TEP>
 A:Cross-references: GB:M33753
 A:Note: the authors translated the codon GGC for residue 1928 as Cys, and TAT for residue
 C:Genetics:
 A:Gene: FlyBase:crb
 A:Cross-references: FlyBase:FBgn0000368
 C:Superfamily: unassigned EGF-related proteins; EGF homology
 C:Keywords: transmembrane protein
 F;352-385/Domain: EGF homology <EGX1>
 F;392-424/Domain: EGF homology <EGF1>
 F;691-722/Domain: EGF homology <EGF>
 F;767-799/Domain: EGF homology <EGF3>
 F;1878-1914/Domain: EGF homology <EGX2>

Query Match 1.9%; Score 9; DB 2; Length 2139;
 Best Local Similarity 100.0%; Pred. No. 1.6;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 48 CECPEGFAG 56
 |||||
 Db 2058 CECPEGFAG 2066

RESULT 11
 A47004
 coagulation factor VIII precursor - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 09-Sep-1994 #sequence_revision 09-Sep-1994 #text_change 18-Jun-1999
 C:Accession: A47004
 R:Elder, B.; Lakich, D.; Gitschier, J.
 Genomics 16, 374-379, 1993
 A:Title: Sequence of the murine factor VIII cDNA.
 A:Reference number: A47004; MUID:93300511
 A:Accession: A47004
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-2319 <ELD>
 A:Cross-references: GB:L05573; NID:g192456; PID:AAA37385.1; PID:g192457
 C:Superfamily: coagulation factor VIII; discoidin I amino-terminal homology; ferroxidase
 F;1-19/Domain: signal sequence #status predicted <SIG>
 F;23-349/Domain: ferroxidase repeat homology <FO1>
 F;402-730/Domain: ferroxidase repeat homology <FO2>
 F;1586-2006/Domain: ferroxidase repeat homology <FO3>

F;2007-2156/Domain: discoidin I amino-terminal homology <DN1>
 F;2160-2313/Domain: discoidin I amino-terminal homology <DN2>

Query Match 1.9%; Score 9; DB 2; Length 2319;
 Best Local Similarity 100.0%; Pred. No. 1.7;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 387 KVTGIITQG 395
 |||||
 Db 2226 KVTGIITQG 2234

RESULT 12
 S42612
 transmembrane protein precursor - zebra fish
 C:Species: Brachydanio rerio (zebra fish)
 C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 20-Sep-1999
 C:Accession: S42612
 R:Bierkamp, C.; Campos-Ortega, J.A.
 Mech. Dev. 43, 87-100, 1993
 A:Title: A zebrafish homologue of the Drosophila neurogenic gene Notch and its patter
 A:Reference number: S42612; MUID:94128602
 A:Accession: S42612
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-2437 <BIE>
 A:Cross-references: EMBL:X69088; NID:g433866; PID:CAA48831.1; PID:g433867
 C:Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homol
 F;755-786/Domain: EGF homology <EGF1>
 F;1023-1054/Domain: EGF homology <EGF2>
 F;1185-1216/Domain: EGF homology <EGF2>
 F;1915-1947/Domain: ankyrin repeat homology <AN1>
 F;1948-1980/Domain: ankyrin repeat homology <AN2>
 F;1982-2014/Domain: ankyrin repeat homology <AN3>
 F;2015-2047/Domain: ankyrin repeat homology <AN4>
 F;2048-2080/Domain: ankyrin repeat homology <AN5>

Query Match 1.9%; Score 9; DB 2; Length 2437;
 Best Local Similarity 100.0%; Pred. No. 1.8;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 NPCHNGTGC 89
 |||||
 Db 683 NPCHNGTGC 691

RESULT 13
 S18188
 notch protein homolog - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 19-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 13-Aug-1999
 C:Accession: S18188
 R:Weinmaster, G.; Roberts, V.J.; Lemke, G.
 Development 113, 199-205, 1991
 A:Title: A homolog of Drosophila Notch expressed during mammalian development.
 A:Reference number: S18188; MUID:92111383
 A:Accession: S18188
 A:Molecule type: mRNA
 A:Residues: 1-2531 <WEI>
 A:Cross-references: EMBL:X57405; NID:g57634; PID:g57635
 C:Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homol
 F;987-1018/Domain: EGF homology <EGF1>
 F;1025-1056/Domain: EGF homology <EGF2>
 F;1233-1264/Domain: EGF homology <EGF2>
 F;1917-1949/Domain: ankyrin repeat homology <AN1>
 F;1950-1982/Domain: ankyrin repeat homology <AN2>
 F;1984-2016/Domain: ankyrin repeat homology <AN3>
 F;2017-2049/Domain: ankyrin repeat homology <AN4>
 F;2050-2082/Domain: ankyrin repeat homology <AN5>

Query Match 1.9%; Score 9; DB 2; Length 2531;
 Best Local Similarity 100.0%; Pred. No. 1.9;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 82 PCHNGTCE 90
 |||||||
 Db 686 PCHNGTCE 694

RESULT 14
 A46019
 Notch-1 protein - mouse
 N:Alternate names: notch protein
 C:Species: Mus musculus (house mouse)
 C:Date: 22-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 20-Sep-1999
 C:Accession: A46019; S25144
 R:del Amo, F.F.; Gendron-Maguire, M.; Swiatek, P.J.; Jenkins, N.A.; Copeland, N.G.; Grid
 Genomics 15, 259-264, 1993
 A:Title: Cloning, analysis, and chromosomal localization of Notch-1, a mouse homolog of
 A:Reference number: A46019; MUID:93194170
 A:Accession: A46019
 A:Status: not compared with conceptual translation
 A:Molecule type: nucleic acid
 A:Residues: 1-2531 <DBL>
 A:Cross-references: GB:Z11886; GB:S47228; NID:g288502; PIDN:CAA77941.1; PID:g288503
 A:Note: sequence extracted from NCBI backbone (NCBIP:127318)
 R:Francisco del Amo, F.; Smith, D.E.; Swiatek, P.J.; Gendron-Maguire, M.; Greenspan, R.J.;
 submitted to the EMBL Data Library, April 1992
 A:Description: Expression pattern of Notch, a mouse homolog of Drosophila Notch, suggest
 A:Reference number: S25144
 A:Accession: S25144
 A:Molecule type: mRNA
 A:Residues: 1551-2108, 'O', 2110-2114, 'ALP', 2118-2170 <FRA>
 A:Cross-references: EMBL:Z11886
 C:Genetics:
 A:Gene: notch-1
 A:Map position: 2

A:Note: proximal region of chromosome 2
 C:Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homology
 F:106-138/Domain: EGF homology <EGF1>
 F:144-175/Domain: EGF homology <EG01>
 F:222-254/Domain: EGF homology <EGP2>
 F:261-292/Domain: EGF homology <EG02>
 F:339-370/Domain: EGF homology <EG03>
 F:416-449/Domain: EGF homology <EGP3>
 F:456-487/Domain: EGF homology <EG04>
 F:494-525/Domain: EGF homology <EG05>
 F:532-563/Domain: EGF homology <EG06>
 F:607-638/Domain: EGF homology <EG07>
 F:682-713/Domain: EGF homology <EG08>
 F:757-788/Domain: EGF homology <EG09>
 F:795-826/Domain: EGF homology <EG10>
 F:873-904/Domain: EGF homology <EG11>
 F:911-942/Domain: EGF homology <EG12>
 F:949-980/Domain: EGF homology <EG13>
 F:987-1018/Domain: EGF homology <EG14>
 F:1025-1056/Domain: EGF homology <EG15>
 F:1063-1094/Domain: EGF homology <EG16>
 F:1149-1180/Domain: EGF homology <EG17>
 F:1187-1218/Domain: EGF homology <EG18>
 F:1233-1264/Domain: EGF homology <EGF4>
 F:1352-1383/Domain: EGF homology <EG19>
 F:1391-1425/Domain: EGF homology <EGF>
 F:1917-1948/Domain: ankyrin repeat homology <AN1>
 F:1949-1981/Domain: ankyrin repeat homology <AN2>
 F:1983-2015/Domain: ankyrin repeat homology <AN3>
 F:2016-2048/Domain: ankyrin repeat homology <AN4>
 F:2049-2081/Domain: ankyrin repeat homology <AN5>

Query Match 1.9%; Score 9; DB 2; Length 2531;
 Best Local Similarity 100.0%; Pred. No. 1.9;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 82 PCHNGTCE 90
 |||||||
 Db 686 PCHNGTCE 694

RESULT 15
 F75371
 hypothetical protein - Deinococcus radiodurans (strain R1)
 C:Species: Deinococcus radiodurans
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
 C:Accession: F75371
 R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
 Science 286, 1571-1577, 1999
 A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
 A:Reference number: A75250; MUID:20036896
 A:Accession: F75371
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-63 <WHI>
 A:Cross-references: GB:AE002007; GB:AE000513; NID:g6459402; PIDN:AAF11201.1; PID:g645
 A:Experimental source: strain R1
 C:Genetics:
 A:Gene: DR1640
 A:Map position: 1

Query Match 1.7%; Score 8; DB 2; Length 63;
 Best Local Similarity 100.0%; Pred. No. 0.75;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 AWLLVGLS 14
 |||||||
 Db 44 AWLLVGLS 51

RESULT 16
 E75609
 amino acid ABC transporter, permease protein - Deinococcus radiodurans (strain R1)
 C:Species: Deinococcus radiodurans
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
 C:Accession: E75609
 R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
 Science 286, 1571-1577, 1999
 A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
 A:Reference number: A75250; MUID:20036896
 A:Accession: E75609
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-247 <WHI>
 A:Cross-references: GB:AE001862; GB:AE001825; NID:g6460468; PIDN:AAF12222.1; PID:g646
 A:Experimental source: strain R1
 C:Genetics:
 A:Gene: DRA0138
 A:Map position: 2
 C:Superfamily: glycine betaine/carnitine/choline ABC transporter

Query Match 1.7%; Score 8; DB 2; Length 247;
 Best Local Similarity 100.0%; Pred. No. 2.6;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 LLVGLSLG 16
 |||||||
 Db 67 LLVGLSLG 74

RESULT 17
 T46256

```

brevican - human (fragment)
N:Alternate names: protein DKFzp761L191.1
C:Species: Homo sapiens (man)
C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 02-Sep-2000
C:Accession: T46256
R:Ottewaelder, B.; Obermaier, B.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, January 2000
A:Reference number: 223031
A:Accession: T46256
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-330 <AAA>
A:Cross-references: EMBL:AL137504
A:Experimental source: adult amygdala; clone DKFzp761L191
C:Genetics:
C:Superfamily: aggregran; C-type lectin homology; complement factor H repeat homology; EG

```

```

Query Match 1.7%; Score 8; DB 2; Length 330;
Best Local Similarity 100.0%; Pred. No. 3.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 82 PCHNGGTC 89
    |||||
DB 73 PCHNGGTC 80

RESULT 18
T48516
probable oligopeptide transporter protein - Arabidopsis thaliana
N:Alternate names: protein F15N18.160
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C:Accession: T48516
R:Bevan, M.; Hilbert, H.; Braun, M.; Holzer, E.; Brandt, A.; Duesterhoeft, A.; Bancroft,
submitted to the Protein Sequence Database, April 2000
A:Reference number: 224490
A:Accession: T48516
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-481 <BEV>
A:Cross-references: EMBL:AL163815
A:Experimental source: cultivar Columbia; BAC clone F15N18
C:Genetics:
A:Map position: 5
A:Introns: 25/1; 97/3; 233/1; 364/3
A:Note: F15N18.160

Query Match 1.7%; Score 8; DB 2; Length 481;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 WLLVGLSL 15
    |||||
DB 456 WLLVGLSL 463

RESULT 19
A31246
neurogenic protein Delta precursor - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 11-Jan-2000
C:Accession: A31246
R:Kopczynski, C.C.; Altan, A.K.; Fechtel, K.; Kooh, P.J.; Muskavitch, M.A.T.
Genes Dev. 2, 1723-1735, 1988
A:Title: Delta, a Drosophila neurogenic gene, is transcriptionally complex and encodes a
A:Reference number: A31246; MUID:89196890
A:Accession: A31246
A:Molecule type: mRNA
A:Residues: 1-832 <KOP>
A:Cross-references: GB:Y00222

```

```

C:Genetics:
A:Gene: FlyBase:Dl
A:Cross-references: FlyBase:FBgn0000463
C:Superfamily: unassigned EGF-related proteins; EGF homology
F:293-338/Domain: EGF homology <EGF1>
F:422-450/Domain: EGF homology <EGF1>
F:457-488/Domain: EGF homology <EGF>
F:533-564/Domain: EGF homology <EGF3>

Query Match 1.7%; Score 8; DB 2; Length 832;
Best Local Similarity 100.0%; Pred. No. 7.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 82 PCHNGGTC 89
    |||||
DB 537 PCHNGGTC 544

RESULT 20
S19087
gene Delta protein precursor - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 11-Jan-2000
C:Accession: S19087
R:Muskavitch, M.A.T.
submitted to the EMBL Data Library, June 1991
A:Reference number: S19087
A:Accession: S19087
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-833 <MUS>
A:Cross-references: EMBL:Y00222
C:Genetics:
A:Gene: FlyBase:Dl
A:Cross-references: FlyBase:FBgn0000463
C:Superfamily: unassigned EGF-related proteins; EGF homology
F:335-371/Domain: EGF homology <EGF1>
F:378-415/Domain: EGF homology <EGF1>
F:457-488/Domain: EGF homology <EGF>
F:533-564/Domain: EGF homology <EGF3>

Query Match 1.7%; Score 8; DB 2; Length 833;
Best Local Similarity 100.0%; Pred. No. 7.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 82 PCHNGGTC 89
    |||||
DB 537 PCHNGGTC 544

RESULT 21
S00670
neurogenic repetitive locus delta protein precursor - fruit fly (Drosophila melanogaster)
N:Alternate names: gene Dl protein
C:Species: Drosophila melanogaster
C:Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 21-Jul-2000
C:Accession: S00670; A26637
R:Vaessin, H.; Bremer, K.A.; Knust, E.; Campos-Ortega, J.A.
EMBO J. 6, 3431-3440, 1987
A:Title: The neurogenic gene Delta of Drosophila melanogaster is expressed in neuroge
A:Reference number: S00670
A:Accession: S00670
A:Molecule type: mRNA
A:Residues: 1-880 <VAE>
A:Cross-references: EMBL:X06289; NID:g7852; PID:g7853
R:Knust, E.; Dietrich, U.; Tepass, U.; Bremer, K.A.; Weigelt, D.; Vaessin, H.; Campos-
EMBO J. 6, 761-766, 1987
A:Title: EGF homologous sequences encoded in the genome of Drosophila melanogaster, a
A:Reference number: A91081; MUID:87218537
A:Accession: A26637
A:Molecule type: mRNA

```

A:Residues: 422-436, 'ET', 439-458, 'A', 460-489, 'T', 491-621 <KNU>
 A:Cross-references: GB:X05140; NID:g7851; PIDN:CAA28786.1; PID:g929563
 C:Genetics:
 A:Gene: Delta; D1
 A:Cross-references: FlyBase:FBgn0000463
 C:Superfamily: unassigned EGF-related proteins; EGF homology
 C:Keywords: transmembrane protein
 F:1-18/Domain: signal sequence #status predicted <SIG>
 F:19-88/Product: neurogenic repetitive locus delta protein #status predicted <MAT>
 F:457-488/Domain: EGF homology <EGF1>
 F:533-564/Domain: EGF homology <EGF2>

Query Match 1.7%; Score 8; DB 2; Length 880;
 Best Local Similarity 100.0%; Pred. No. 8.3;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 82 PCHNGGTC 89

Db 537 PCHNGGTC 544

RESULT 22
 S49126
 brevicin precursor - rat
 N:Alternate names: aggrecan-like protein
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 01-Feb-1995 #sequence_revision 12-May-1995 #text_change 26-May-2000
 C:Accession: S49126; I55457
 R:Seidenbecher, C.I.; Langnese, K.; Wax, H.; Seidel, B.; Garner, C.C.; Gundelfinger, E.D.
 submitted to the EMBL Data Library, June 1994
 A:Description: Molecular cloning of a new member of the aggrecan/versican family of proteoglycans.
 A:Reference number: S49126
 A:Accession: S49126
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-883 <SEI>
 A:Cross-references: EMBL:X79881; NID:g509396; PIDN:CAA56255.1; PID:g509397
 R:Seidenbecher, I.C.; Richter, K.; Rauch, U.; Faessler, R.; Garner, C.C.; Gundelfinger, E.D.
 J. Biol. Chem. 270, 27206-27212, 1995
 A:Title: Brevican, a Chondroitin Sulfate Proteoglycan of Rat Brain, Occurs as Secreted and Membrane-associated Forms.
 A:Reference number: I55457; MUID:96070828
 A:Accession: I55457
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-883 <RES>
 A:Cross-references: EMBL:X79881; NID:g509396; PIDN:CAA56255.1; PID:g509397
 R:Seidenbecher, I.C.; Richter, K.; Rauch, U.; Faessler, R.; Garner, C.C.; Gundelfinger, E.D.
 J. Biol. Chem. 270, 27206-27212, 1995
 A:Comment: For an alternative splice form, see PIR:A53908.
 C:Superfamily: aggrecan; C-type lectin homology; complement factor H repeat homology; EGF homology
 C:Keywords: alternative splicing
 F:49-138/Domain: immunoglobulin homology <IMM>
 F:173-250/Domain: link protein repeat homology <LNK1>
 F:271-352/Domain: link protein repeat homology <LNK2>
 F:626-657/Domain: EGF homology <EGF>
 F:664-784/Domain: C-type lectin homology <LCH>
 F:791-847/Domain: complement factor H repeat homology <FHD>

Query Match 1.7%; Score 8; DB 2; Length 883;
 Best Local Similarity 100.0%; Pred. No. 8.3;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 82 PCHNGGTC 89

Db 630 PCHNGGTC 637

RESULT 23
 S57653
 brevicin precursor - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 19-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 13-Aug-1999
 C:Accession: S57653

R:Rauch, U.; Forsberg, N.; Kulbe, G.; Arnold-Ammer, I.; Faessler, R.
 submitted to the EMBL Data Library, May 1995
 A:Description: Cloning and sequence of mouse neurocan and brevican and their differences.
 A:Reference number: S57653
 A:Accession: S57653
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-883 <RAU>
 A:Cross-references: EMBL:X87096; NID:g886889; PIDN:CAA60575.1; PID:g886890
 C:Superfamily: aggrecan; C-type lectin homology; complement factor H repeat homology;
 F:49-138/Domain: immunoglobulin homology <IMM>
 F:173-250/Domain: link protein repeat homology <LNK1>
 F:271-352/Domain: link protein repeat homology <LNK2>
 F:626-657/Domain: EGF homology <EGF>
 F:664-784/Domain: C-type lectin homology <LCH>
 F:791-847/Domain: complement factor H repeat homology <FHD>

Query Match 1.7%; Score 8; DB 2; Length 883;
 Best Local Similarity 100.0%; Pred. No. 8.3;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 82 PCHNGGTC 89

Db 630 PCHNGGTC 637

RESULT 24
 A54423
 brevicin precursor - bovine
 C:Species: Bos primigenius taurus (cattle)
 C:Date: 09-Sep-1994 #sequence_revision 09-Sep-1994 #text_change 21-Jan-2000
 C:Accession: A54423; S41914
 R:Yamada, H.; Watanabe, K.; Shimonaka, M.; Yamaguchi, Y.
 J. Biol. Chem. 269, 10119-10126, 1994
 A:Title: Molecular cloning of brevicin, a novel brain proteoglycan of the aggrecan/versican family.
 A:Reference number: A54423; MUID:94193597
 A:Accession: A54423
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-912 <YAM>
 A:Cross-references: GB:X75887; NID:g452820; PIDN:CAA53481.1; PID:g452821
 C:Superfamily: aggrecan; C-type lectin homology; complement factor H repeat homology;
 F:50-139/Domain: immunoglobulin homology <IMM>
 F:174-251/Domain: link protein repeat homology <LNK1>
 F:272-353/Domain: link protein repeat homology <LNK2>
 F:651-682/Domain: EGF homology <EGF>
 F:689-809/Domain: C-type lectin homology <LCH>
 F:816-872/Domain: complement factor H repeat homology <FHD>

Query Match 1.7%; Score 8; DB 2; Length 912;
 Best Local Similarity 100.0%; Pred. No. 8.5;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 82 PCHNGGTC 89

Db 655 PCHNGGTC 662

RESULT 25
 T21133
 hypothetical protein F20B10.1 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T21133
 R:Percy, C.
 submitted to the EMBL Data Library, February 1996
 A:Reference number: Z19380
 A:Accession: T21133
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1193 <WIL>

A;Cross-references: EMBL:269636; PIDN:CAA93465.1; GSPDB:GN00022; CESP:F20B10.1
A;Experimental source: clone F20B10
C;Genetics:
A;Gene: CESP:F20B10.1
A;Map position: 4
A;Introns: 18/1; 74/3; 135/1; 222/3; 258/3; 279/1; 379/1; 496/1; 791/2; 867/1; 969/1; 10

Query Match 1.7%; Score 8; DB 2; Length 1193;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 PNCENG 35
Db 463 PNCENG 470

Search completed: May 23, 2001, 06:22:17
Job time: 236 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 06:21:46 ; Search time 17.36 seconds

(without alignments)

947.156 Million cell updates/sec

Title: US-09-237-981-10

Perfect score: 480

Sequence: 1 MKHLVAALLVGLSLGVPQF.....WSWYGRITLRSLLGCAEE 480

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 93435 seqs, 34255486 residues

Word size : 0

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 50 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	10	2.1	409	1 MFGM_PIG	P79385 sus scrofa
2	10	2.1	427	1 MFGM_RAT	P70490 rattus norv
3	10	2.1	914	1 NRPL_CHICK	P79795 gallus gall
4	10	2.1	2444	1 NTC1_HUMAN	P46531 homo sapien
5	9	1.9	387	1 MFGM_HUMAN	Q08431 homo sapien
6	9	1.9	427	1 MFGM_BOVIN	Q95114 bos taurus
7	9	1.9	463	1 MFGM_MOUSE	P21956 mus musculu
8	9	1.9	922	1 NRPL_RAT	Q9QWJ9 rattus norv
9	9	1.9	923	1 NRPL_HUMAN	O14786 homo sapien
10	9	1.9	923	1 NRPL_MOUSE	P97333 mus musculu
11	9	1.9	1408	1 SERR_DROME	P18168 drosophila
12	9	1.9	2139	1 CRB_DROME	P10040 drosophila
13	9	1.9	2319	1 FAB_MOUSE	Q06194 mus musculu
14	9	1.9	2437	1 NOTC_BRARE	P46530 brachydanio
15	9	1.9	2531	1 NTC1_MOUSE	Q01705 mus musculu
16	8	1.9	2531	1 NTC1_RAT	Q07008 rattus norv
17	8	1.7	833	1 DL_DROME	P10041 drosophila
18	8	1.7	883	1 PGCB_MOUSE	Q61361 mus musculu
19	8	1.7	883	1 PGCB_RAT	P55068 rattus norv
20	8	1.7	912	1 PGCB_BOVIN	Q28062 bos taurus
21	8	1.7	2524	1 NOTC_XENLA	P21783 xenopus lae
22	8	1.7	2670	1 YAO5_SCHPO	Q10105 schizosacch
23	7	1.5	106	1 SH_RAT	P55248 rattus norv
24	7	1.5	110	1 CYOD_PSEPU	Q9WWR4 pseudomonas
25	7	1.5	149	1 CRAA_RANTE	P22508 rana ran
26	7	1.5	167	1 CRAA_RANES	P02507 rana escul
27	7	1.5	173	1 CRAA_ALIMI	P06904 alligator m
28	7	1.5	173	1 CRAA_CHICK	P02504 gallus gall
29	7	1.5	173	1 CRAA_HUMAN	P02489 homo sapien
30	7	1.5	173	1 CRAA_RANCA	Q91311 rana catesb
31	7	1.5	173	1 CRAA_RHEAM	P02505 rhea ameri
32	7	1.5	173	1 CRAA_TUPT	P02506 tupinambis
33	7	1.5	259	1 YL23_YEAST	P53769 saccharomyc

ALIGNMENTS

RESULT 1

MFGM_PIG	ID	MFGM_PIG	STANDARD;	PRT;	409 AA.
AC	P79385;				
DT	01-NOV-1997 (Rel. 35, Created)				
DT	15-JUL-1998 (Rel. 36, Last sequence update)				
DT	30-MAY-2000 (Rel. 39, Last annotation update)				
DE	LACTADHERIN (MILK FAT GLOBULE-EGF FACTOR 8) (MFG-EB) (MFGM) (SPERM				
DE	SURFACE PROTEIN SP47) (PP47).				
GN	MFGM.				
OS	Sus scrofa (Pig).				
OC	Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.				
OX	NCBI_TaxID=9823;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Testis;				
RA	Ensslin M.A.;				
RL	Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.				
CC	-!- FUNCTION: MAY BE INVOLVED IN PHOSPHOLIPID BINDING. ZONA PELLUCIDA-BINDING PROTEIN.				
CC	-!- SUBCELLULAR LOCATION: PERIPHERAL MEMBRANE PROTEIN				
CC	-!- TISSUE SPECIFICITY: MAMMARY EPITHELIAL CELL SURFACES AND SPERMATOZOAN				
CC	-!- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.				
CC	-!- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch)				
CC	EMBL; Y11683; CA472379.1; -				
DR	HSSP; P00740; I1XA.				
DR	InterPro; IPR000421; -				
DR	InterPro; IPR000561; -				
DR	Pfam; PF00008; EGF; 2.				
DR	Pfam; PF00754; F5_F8_typeC; 2.				
DR	PROSITE; PS00022; EGF_1; 2.				
DR	PROSITE; PS01186; EGF_2; 2.				
DR	PROSITE; PS01285; FA58C_1; 2.				
DR	PROSITE; PS01286; FA58C_2; 2.				
KW	Glycoprotein; Repeat; EGF-like domain.				
FT	DOMAIN 2 41				
FT	DOMAIN 44 88				
FT	EGF-LIKE 1.				
FT	EGF-LIKE 2.				
FT	F5/8 TYPE C.				
FT	DOMAIN 91 247				
FT	DOMAIN 252 409				
FT	F5/8 TYPE C.				
FT	CELL ATTACHMENT SITE (POTENTIAL).				
FT	SITE 67 69				

P27685 dictyosteli
Q9zd77 rickettsia
P32166 escherichia
P31997 homo sapien
Q99678 homo sapien
Q07172 streptococ
P08369 escherichia
P00750 homo sapien
Q9VYV4 drosophila
P07304 homo sapien
P15306 mus musculu
P79755 fugu rubrip
P54421 rhizobium s
Q05996 homo sapien
Q05586 homo sapien
P35438 mus musculu
P35439 rattus norv

34 7 1.5 265 1 RS2_DICDI
35 7 1.5 299 1 HEM3_RICPR
36 7 1.5 308 1 MENA_ECOLI
37 7 1.5 349 1 CGM6_HUMAN
38 7 1.5 358 1 GPRK_HUMAN
39 7 1.5 402 1 UDG_STRPY
40 7 1.5 450 1 CRED_ECOLI
41 7 1.5 562 1 TPA_HUMAN
42 7 1.5 574 1 C4GF_DROME
43 7 1.5 575 1 TRBM_MOUSE
44 7 1.5 577 1 TRBM_MOUSE
45 7 1.5 586 1 CO9_FUGRU
46 7 1.5 640 1 TRAG_RHISN
47 7 1.5 745 1 ZP2_HUMAN
48 7 1.5 938 1 NM21_HUMAN
49 7 1.5 938 1 NM21_MOUSE
50 7 1.5 938 1 NM21_RAT

```

FT  DISULFID      6  17  BY SIMILARITY.
FT  DISULFID     11  29  BY SIMILARITY.
FT  DISULFID     31  40  BY SIMILARITY.
FT  DISULFID     91  247 BY SIMILARITY.
FT  DISULFID    234  238 BY SIMILARITY.
FT  DISULFID    252  409 BY SIMILARITY.
FT  DISULFID     41  41  N-LINKED (GLCNAC. . .) (POTENTIAL).
FT  CARBOHYD     372  372 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ  SEQUENCE    409 AA; 45725 MW; B0C07AF80029927A CRC64;

Query Match      2.1%; Score 10; DB 1; Length 409;
Best Local Similarity 100.0%; Pred. No. 0.02;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy  310 ELLGCELSGC 319
    |||||
Db  243 ELLGCELSGC 252

RESULT 2
ID  MFGM_RAT      STANDARD;      PRT;      427 AA.
AC  P70490;
DT  01-NOV-1997 (Rel. 35, Created)
DT  01-NOV-1997 (Rel. 35, Last sequence update)
DT  30-MAY-2000 (Rel. 39, Last annotation update)
DE  LACTADHERIN PRECURSOR (MILK FAT GLOBULE-EGF FACTOR 8) (MFG-E8) (O-
DE  ACETYL GD3 GANGLIOSIDE SYNTHASE) (AGS) (MFGM).
GN  MFG-E8 OR AGS.
OS  Rattus norvegicus (Rat).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX  NCBI_TaxID=10116;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  TISSUE=Brain;
RA  MEDLINE=96374422; PubMed=8780713;
RX  Oqura K., Nara K., Watanabe Y., Kohno K., Tai T., Sanai Y.;
RT  "Cloning and expression of cDNA for O-acetylation of GD3
RT  ganglioside."
RL  Biochem. Biophys. Res. Commun. 225:932-938(1996).
CC  -!- FUNCTION: MAY BE INVOLVED IN PHOSPHOLIPID BINDING. SEEMS TO
CC  PARTICIPATE IN THE O-ACETYLATION OF GD3 GANGLIOSIDE SIALIC ACID.
CC  -!- SUBCELLULAR LOCATION: PERIPHERAL MEMBRANE PROTEIN.
CC  -!- TISSUE SPECIFICITY: SPLEEN, LUNG, HEART, BRAIN AND MUSCLE.
CC  -!- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
CC  -!- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
CC  EMBL; D84068; BAA12210.1; -
CC  HSSP; P00740; 11XA.
CC  InterPro; IPR000421; -
CC  InterPro; IPR000561; -
CC  InterPro; IPR001438; -
CC  Pfam; PF00008; EGF; 2.
CC  Pfam; PF00754; F5_F8_type_C; 2.
CC  PRINTS; PR00010; EGFBL00D.
CC  PROSITE; PS00022; EGF_1; 2.
CC  PROSITE; PS01186; EGF_2; 2.
CC  PROSITE; PS01285; FAS8C_1; 2.
CC  PROSITE; PS01286; FAS8C_2; 2.
CC  Signal; Glycoprotein; Repeat; EGF-like domain; Milk.
CC  SIGNAL 1 22 POTENTIAL.
CC  CHAIN 23 427 LACTADHERIN.
CC  DOMAIN 24 61 EGF-LIKE 1.

```

```

FT  DOMAIN      64  108 EGF-LIKE 2.
FT  DOMAIN     111  267 F5/8 TYPE C 1.
FT  DOMAIN     272  427 F5/8 TYPE C 2.
FT  DISULFID     28  39  BY SIMILARITY.
FT  DISULFID     33  49  BY SIMILARITY.
FT  DISULFID     51  60  BY SIMILARITY.
FT  DISULFID     68  79  BY SIMILARITY.
FT  DISULFID     73  96  BY SIMILARITY.
FT  DISULFID     98  107 BY SIMILARITY.
FT  DISULFID    111  267 BY SIMILARITY.
FT  DISULFID    254  258 BY SIMILARITY.
FT  DISULFID    272  427 BY SIMILARITY.
FT  SITE        87  89  CELL ATTACHMENT SITE (POTENTIAL).
FT  CARBOHYD     61  61  N-LINKED (GLCNAC. . .) (POTENTIAL).
FT  CARBOHYD    230  230 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT  CARBOHYD    280  280 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT  CARBOHYD    390  390 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ  SEQUENCE    427 AA; 47413 MW; EA8C8631F3EE6047 CRC64;

```

```

Query Match      2.1%; Score 10; DB 1; Length 427;
Best Local Similarity 100.0%; Pred. No. 0.02;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy  387 KVTGIITQGA 396
    |||||
Db  340 KVTGIITQGA 349

RESULT 3
ID  NRPL_CHICK      STANDARD;      PRT;      914 AA.
AC  P79795;
DT  01-NOV-1997 (Rel. 35, Created)
DT  01-NOV-1997 (Rel. 35, Last sequence update)
DT  01-OCT-2000 (Rel. 40, Last annotation update)
DE  NEUROFILIN-1 PRECURSOR (A5 PROTEIN).
DE  NRPL OR NRP.
GN  Gallus gallus (Chicken).
OS  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC  Gallus.
OX  NCBI_TaxID=9031;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=WHITE LEGHORN; TISSUE=Embryonic brain;
RX  MEDLINE=95324761; PubMed=7601310;
RX  Takagi S., Kasuya Y., Shimizu M., Matsuura T., Tsuboi M., Kawakami A.,
RX  Fujisawa H.;
RT  "Expression of a cell adhesion molecule, neuropilin, in the
RT  developing chick nervous system."
RL  Dev. Biol. 170:207-222(1995).
CC  -!- FUNCTION: RECEPTOR INVOLVED IN THE DEVELOPMENT OF THE
CC  CARDIOVASCULAR SYSTEM, IN ANGIOGENESIS, IN THE FORMATION OF
CC  CERTAIN NEURONAL CIRCUITS AND IN ORGANOGENESIS OUTSIDE THE NERVOUS
CC  SYSTEM. IT MEDIATES THE CHEMOREPULSANT ACTIVITY OF SEMAPHORINS (BY
CC  SIMILARITY). SEEMS TO HAVE CALCIUM-INDEPENDENT CELL ADHESION
CC  PROPERTIES.
CC  -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC  -!- TISSUE SPECIFICITY: DEVELOPING NERVOUS SYSTEM; OPTIC TECTUM
CC  (LAYERS D AND E OF SGFS); AMACRINE CELLS OF RETINA, NEURITES OF
CC  DORSAL ROOT GANGLIA. ALSO EXPRESSED IN NONNEURONAL CELLS, E.G.
CC  BLOOD VESSELS IN THE ENTIRE EMBRYO.
CC  -!- SIMILARITY: BELONGS TO THE NEUROFILIN FAMILY.
CC  -!- SIMILARITY: CONTAINS 2 CUB DOMAINS.
CC  -!- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.
CC  -!- SIMILARITY: CONTAINS 1 NAM DOMAIN.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----

```

CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

```

CC EMBL: D45416; BAA08256.1; -
CC InterPro: IPR000421; -
CC InterPro: IPR000859; -
CC InterPro: IPR000998; -
CC Pfam: PF00431; CUB; 2.
CC Pfam: PF00754; F5_F8_type_C; 2.
CC Pfam: PF00629; MAM; 1.
CC PRINTS: PR00020; MAMDOMAIN.
CC PROSITE: PS01180; CUB; 2.
CC PROSITE: PS01285; FA58C_1; 2.
CC PROSITE: PS01286; FA58C_2; 2.
CC PROSITE: PS00740; MAM_1; 1.
CC PROSITE: PS00600; MAM_2; 1.
CC Transmembrane; Glycoprotein; Neurone; Signal; Repeat; Receptor;
KW Cell adhesion.
FT SIGNAL 1 18
FT CHAIN 19 914
FT DOMAIN 20 847
FT TRANSMEM 848 870
FT DOMAIN 871 914
FT DOMAIN 25 139
FT DOMAIN 145 263
FT DOMAIN 273 422
FT DOMAIN 429 581
FT DOMAIN 636 801
FT DISULFID 25 52
FT DISULFID 80 102
FT DISULFID 145 171
FT DISULFID 204 226
FT DISULFID 273 422
FT DISULFID 429 581
SQ SEQUENCE 914 AA; 102480 MW; DD2EE6D6FCBB68C CRC64;

Query Match 2.1%; Score 10; DB 1; Length 914;
Best Local Similarity 100.0%; Pred. No. 0.04;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 307 LRMEELGGCEL 316
Db 574 LRMEELGGCEL 583
|||||

RESULT 4
NTCL_HUMAN
ID NTCL_HUMAN STANDARD; PRT; 2444 AA.
AC P46531;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE NEUROGENIC LOCUS NOTCH PROTEIN HOMOLOG 1 PRECURSOR (TRANSLOCATION-
DE ASSOCIATED NOTCH PROTEIN TAN-1) (FRAGMENT).
GN NOTCH1 OR TAN1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91347367; PubMed=1831692;
RA Ellisen L.W., Bird J., West D.C., Soreng A.L., Reynolds T.C.,
RA Smith S.D., Sklar J.;
RT "TAN-1, the human homolog of the Drosophila notch gene, is broken by
RT chromosomal translocations in T lymphoblastic neoplasms.";
RL Cell 66:649-661(1991).
CC -|- FUNCTION: MAY BE IMPORTANT FOR NORMAL LYMPHOCYTE FUNCTION. IN
CC ALTERED FORM, MAY CONTRIBUTE TO TRANSFORMATION OR PROGRESSION
CC IN SOME T-CELL NEOPLASMS.
CC -|- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -|- TISSUE SPECIFICITY: IN FETAL TISSUES MOST ABUNDANT IN SPLEEN,

```

```

CC BRAIN STEM AND LUNG. ALSO PRESENT IN MOST ADULT TISSUES WHERE IT
CC IS FOUND MAINLY IN LYMPHOID TISSUES.
CC -|- SIMILARITY: HIGH, WITH OTHER NOTCH-TYPE PROTEINS.
CC -|- SIMILARITY: CONTAINS 36 EGF-LIKE DOMAINS.
CC -|- SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
CC -|- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).


EMBL: M73980; AAA60614.1; -
HSP: P00740; LIXA.
MIM: 190198; -
InterPro: IPR000152; -
InterPro: IPR000561; -
InterPro: IPR000800; -
InterPro: IPR001881; -
InterPro: IPR002110; -
Pfam: PF00008; EGF; 36.
Pfam: PF00023; ank; 6.
Pfam: PF00066; notch; 3.
PROSITE: PS50088; ANK_REPEAT; 4.
PROSITE: PS50297; ANK_REPEAT_REGION; 1.
PROSITE: PS00010; ASX_HYDROXYL; 20.
PROSITE: PS00022; EGF_1; 34.
PROSITE: PS01186; EGF_2; 26.
PROSITE: PS01187; EGF_CA; 18.
KW Differentiation; Neurogenesis; Repeat; ANK repeat; EGF-like domain;
KW Transmembrane; Signal; Glycoprotein.
FT SIGNAL 1 18
FT CHAIN 19 >2444
FT DOMAIN 19 1736
FT TRANSMEM 1737 1757
FT DOMAIN 1758 >2444
FT DOMAIN 20 58
FT DOMAIN 59 99
FT DOMAIN 102 139
FT DOMAIN 140 176
FT DOMAIN 178 216
FT DOMAIN 218 255
FT DOMAIN 257 293
FT DOMAIN 295 333
FT DOMAIN 335 371
FT DOMAIN 372 410
FT DOMAIN 412 450
FT DOMAIN 452 488
FT DOMAIN 490 526
FT DOMAIN 528 564
FT DOMAIN 566 601
FT DOMAIN 603 639
FT DOMAIN 641 676
FT DOMAIN 678 714
FT DOMAIN 716 751
FT DOMAIN 753 789
FT DOMAIN 791 827
FT DOMAIN 829 868
FT DOMAIN 870 906
FT DOMAIN 908 944
FT DOMAIN 946 982
FT DOMAIN 984 1020
FT DOMAIN 1022 1058
FT DOMAIN 1060 1096
FT DOMAIN 1098 1144
FT DOMAIN 1146 1182
FT DOMAIN 1184 1220
FT DOMAIN 1222 1266
FT DOMAIN 1268 1306
FT DOMAIN 1308 1347
FT DOMAIN 1347 1387
FT DOMAIN 1387 1423
FT DOMAIN 1423 1459
FT DOMAIN 1459 1495
FT DOMAIN 1495 1531
FT DOMAIN 1531 1567
FT DOMAIN 1567 1603
FT DOMAIN 1603 1639
FT DOMAIN 1639 1675
FT DOMAIN 1675 1711
FT DOMAIN 1711 1747
FT DOMAIN 1747 1783
FT DOMAIN 1783 1819
FT DOMAIN 1819 1855
FT DOMAIN 1855 1891
FT DOMAIN 1891 1927
FT DOMAIN 1927 1963
FT DOMAIN 1963 1999
FT DOMAIN 1999 2035
FT DOMAIN 2035 2071
FT DOMAIN 2071 2107
FT DOMAIN 2107 2143
FT DOMAIN 2143 2179
FT DOMAIN 2179 2215
FT DOMAIN 2215 2251
FT DOMAIN 2251 2287
FT DOMAIN 2287 2323
FT DOMAIN 2323 2359
FT DOMAIN 2359 2395
FT DOMAIN 2395 2431
FT DOMAIN 2431 2467
FT DOMAIN 2467 2503
FT DOMAIN 2503 2539
FT DOMAIN 2539 2575
FT DOMAIN 2575 2611
FT DOMAIN 2611 2647
FT DOMAIN 2647 2683
FT DOMAIN 2683 2719
FT DOMAIN 2719 2755
FT DOMAIN 2755 2791
FT DOMAIN 2791 2827
FT DOMAIN 2827 2863
FT DOMAIN 2863 2899
FT DOMAIN 2899 2935
FT DOMAIN 2935 2971
FT DOMAIN 2971 3007
FT DOMAIN 3007 3043
FT DOMAIN 3043 3079
FT DOMAIN 3079 3115
FT DOMAIN 3115 3151
FT DOMAIN 3151 3187
FT DOMAIN 3187 3223
FT DOMAIN 3223 3259
FT DOMAIN 3259 3295
FT DOMAIN 3295 3331
FT DOMAIN 3331 3367
FT DOMAIN 3367 3403
FT DOMAIN 3403 3439
FT DOMAIN 3439 3475
FT DOMAIN 3475 3511
FT DOMAIN 3511 3547
FT DOMAIN 3547 3583
FT DOMAIN 3583 3619
FT DOMAIN 3619 3655
FT DOMAIN 3655 3691
FT DOMAIN 3691 3727
FT DOMAIN 3727 3763
FT DOMAIN 3763 3799
FT DOMAIN 3799 3835
FT DOMAIN 3835 3871
FT DOMAIN 3871 3907
FT DOMAIN 3907 3943
FT DOMAIN 3943 3979
FT DOMAIN 3979 4015
FT DOMAIN 4015 4051
FT DOMAIN 4051 4087
FT DOMAIN 4087 4123
FT DOMAIN 4123 4159
FT DOMAIN 4159 4195
FT DOMAIN 4195 4231
FT DOMAIN 4231 4267
FT DOMAIN 4267 4303
FT DOMAIN 4303 4339
FT DOMAIN 4339 4375
FT DOMAIN 4375 4411
FT DOMAIN 4411 4447
FT DOMAIN 4447 4483
FT DOMAIN 4483 4519
FT DOMAIN 4519 4555
FT DOMAIN 4555 4591
FT DOMAIN 4591 4627
FT DOMAIN 4627 4663
FT DOMAIN 4663 4699
FT DOMAIN 4699 4735
FT DOMAIN 4735 4771
FT DOMAIN 4771 4807
FT DOMAIN 4807 4843
FT DOMAIN 4843 4879
FT DOMAIN 4879 4915
FT DOMAIN 4915 4951
FT DOMAIN 4951 4987
FT DOMAIN 4987 5023
FT DOMAIN 5023 5059
FT DOMAIN 5059 5095
FT DOMAIN 5095 5131
FT DOMAIN 5131 5167
FT DOMAIN 5167 5203
FT DOMAIN 5203 5239
FT DOMAIN 5239 5275
FT DOMAIN 5275 5311
FT DOMAIN 5311 5347
FT DOMAIN 5347 5383
FT DOMAIN 5383 5419
FT DOMAIN 5419 5455
FT DOMAIN 5455 5491
FT DOMAIN 5491 5527
FT DOMAIN 5527 5563
FT DOMAIN 5563 5599
FT DOMAIN 5599 5635
FT DOMAIN 5635 5671
FT DOMAIN 5671 5707
FT DOMAIN 5707 5743
FT DOMAIN 5743 5779
FT DOMAIN 5779 5815
FT DOMAIN 5815 5851
FT DOMAIN 5851 5887
FT DOMAIN 5887 5923
FT DOMAIN 5923 5959
FT DOMAIN 5959 5995
FT DOMAIN 5995 6031
FT DOMAIN 6031 6067
FT DOMAIN 6067 6103
FT DOMAIN 6103 6139
FT DOMAIN 6139 6175
FT DOMAIN 6175 6211
FT DOMAIN 6211 6247
FT DOMAIN 6247 6283
FT DOMAIN 6283 6319
FT DOMAIN 6319 6355
FT DOMAIN 6355 6391
FT DOMAIN 6391 6427
FT DOMAIN 6427 6463
FT DOMAIN 6463 6499
FT DOMAIN 6499 6535
FT DOMAIN 6535 6571
FT DOMAIN 6571 6607
FT DOMAIN 6607 6643
FT DOMAIN 6643 6679
FT DOMAIN 6679 6715
FT DOMAIN 6715 6751
FT DOMAIN 6751 6787
FT DOMAIN 6787 6823
FT DOMAIN 6823 6859
FT DOMAIN 6859 6895
FT DOMAIN 6895 6931
FT DOMAIN 6931 6967
FT DOMAIN 6967 7003
FT DOMAIN 7003 7039
FT DOMAIN 7039 7075
FT DOMAIN 7075 7111
FT DOMAIN 7111 7147
FT DOMAIN 7147 7183
FT DOMAIN 7183 7219
FT DOMAIN 7219 7255
FT DOMAIN 7255 7291
FT DOMAIN 7291 7327
FT DOMAIN 7327 7363
FT DOMAIN 7363 7399
FT DOMAIN 7399 7435
FT DOMAIN 7435 7471
FT DOMAIN 7471 7507
FT DOMAIN 7507 7543
FT DOMAIN 7543 7579
FT DOMAIN 7579 7615
FT DOMAIN 7615 7651
FT DOMAIN 7651 7687
FT DOMAIN 7687 7723
FT DOMAIN 7723 7759
FT DOMAIN 7759 7795
FT DOMAIN 7795 7831
FT DOMAIN 7831 7867
FT DOMAIN 7867 7903
FT DOMAIN 7903 7939
FT DOMAIN 7939 7975
FT DOMAIN 7975 8011
FT DOMAIN 8011 8047
FT DOMAIN 8047 8083
FT DOMAIN 8083 8119
FT DOMAIN 8119 8155
FT DOMAIN 8155 8191
FT DOMAIN 8191 8227
FT DOMAIN 8227 8263
FT DOMAIN 8263 8299
FT DOMAIN 8299 8335
FT DOMAIN 8335 8371
FT DOMAIN 8371 8407
FT DOMAIN 8407 8443
FT DOMAIN 8443 8479
FT DOMAIN 8479 8515
FT DOMAIN 8515 8551
FT DOMAIN 8551 8587
FT DOMAIN 8587 8623
FT DOMAIN 8623 8659
FT DOMAIN 8659 8695
FT DOMAIN 8695 8731
FT DOMAIN 8731 8767
FT DOMAIN 8767 8803
FT DOMAIN 8803 8839
FT DOMAIN 8839 8875
FT DOMAIN 8875 8911
FT DOMAIN 8911 8947
FT DOMAIN 8947 8983
FT DOMAIN 8983 9019
FT DOMAIN 9019 9055
FT DOMAIN 9055 9091
FT DOMAIN 9091 9127
FT DOMAIN 9127 9163
FT DOMAIN 9163 9199
FT DOMAIN 9199 9235
FT DOMAIN 9235 9271
FT DOMAIN 9271 9307
FT DOMAIN 9307 9343
FT DOMAIN 9343 9379
FT DOMAIN 9379 9415
FT DOMAIN 9415 9451
FT DOMAIN 9451 9487
FT DOMAIN 9487 9523
FT DOMAIN 9523 9559
FT DOMAIN 9559 9595
FT DOMAIN 9595 9631
FT DOMAIN 9631 9667
FT DOMAIN 9667 9703
FT DOMAIN 9703 9739
FT DOMAIN 9739 9775
FT DOMAIN 9775 9811
FT DOMAIN 9811 9847
FT DOMAIN 9847 9883
FT DOMAIN 9883 9919
FT DOMAIN 9919 9955
FT DOMAIN 9955 9991
FT DOMAIN 9991 10027
FT DOMAIN 10027 10063
FT DOMAIN 10063 10099
FT DOMAIN 10099 10135
FT DOMAIN 10135 10171
FT DOMAIN 10171 10207
FT DOMAIN 10207 10243
FT DOMAIN 10243 10279
FT DOMAIN 10279 10315
FT DOMAIN 10315 10351
FT DOMAIN 10351 10387
FT DOMAIN 10387 10423
FT DOMAIN 10423 10459
FT DOMAIN 10459 10495
FT DOMAIN 10495 10531
FT DOMAIN 10531 10567
FT DOMAIN 10567 10603
FT DOMAIN 10603 10639
FT DOMAIN 10639 10675
FT DOMAIN 10675 10711
FT DOMAIN 10711 10747
FT DOMAIN 10747 10783
FT DOMAIN 10783 10819
FT DOMAIN 10819 10855
FT DOMAIN 10855 10891
FT DOMAIN 10891 10927
FT DOMAIN 10927 10963
FT DOMAIN 10963 10999
FT DOMAIN 10999 11035
FT DOMAIN 11035 11071
FT DOMAIN 11071 11107
FT DOMAIN 11107 11143
FT DOMAIN 11143 11179
FT DOMAIN 11179 11215
FT DOMAIN 11215 11251
FT DOMAIN 11251 11287
FT DOMAIN 11287 11323
FT DOMAIN 11323 11359
FT DOMAIN 11359 11395
FT DOMAIN 11395 11431
FT DOMAIN 11431 11467
FT DOMAIN 11467 11503
FT DOMAIN 11503 11539
FT DOMAIN 11539 11575
FT DOMAIN 11575 11611
FT DOMAIN 11611 11647
FT DOMAIN 11647 11683
FT DOMAIN 11683 11719
FT DOMAIN 11719 11755
FT DOMAIN 11755 11791
FT DOMAIN 11791 11827
FT DOMAIN 11827 11863
FT DOMAIN 11863 11899
FT DOMAIN 11899 11935
FT DOMAIN 11935 11971
FT DOMAIN 11971 12007
FT DOMAIN 12007 12043
FT DOMAIN 12043 12079
FT DOMAIN 12079 12115
FT DOMAIN 12115 12151
FT DOMAIN 12151 12187
FT DOMAIN 12187 12223
FT DOMAIN 12223 12259
FT DOMAIN 12259 12295
FT DOMAIN 12295 12331
FT DOMAIN 12331 12367
FT DOMAIN 12367 12403
FT DOMAIN 12403 12439
FT DOMAIN 12439 12475
FT DOMAIN 12475 12511
FT DOMAIN 12511 12547
FT DOMAIN 12547 12583
FT DOMAIN 12583 12619
FT DOMAIN 12619 12655
FT DOMAIN 12655 12691
FT DOMAIN 12691 12727
FT DOMAIN 12727 12763
FT DOMAIN 12763 12799
FT DOMAIN 12799 12835
FT DOMAIN 12835 12871
FT DOMAIN 12871 12907
FT DOMAIN 12907 12943
FT DOMAIN 12943 12979
FT DOMAIN 12979 13015
FT DOMAIN 13015 13051
FT DOMAIN 13051 13087
FT DOMAIN 13087 13123
FT DOMAIN 13123 13159
FT DOMAIN 13159 13195
FT DOMAIN 13195 13231
FT DOMAIN 13231 13267
FT DOMAIN 13267 13303
FT DOMAIN 13303 13339
FT DOMAIN 13339 13375
FT DOMAIN 13375 13411
FT DOMAIN 13411 13447
FT DOMAIN 13447 13483
FT DOMAIN 13483 13519
FT DOMAIN 13519 13555
FT DOMAIN 13555 13591
FT DOMAIN 13591 13627
FT DOMAIN 13627 13663
FT DOMAIN 13663 13699
FT DOMAIN 13699 13735
FT DOMAIN 13735 13771
FT DOMAIN 13771 13807
FT DOMAIN 13807 13843
FT DOMAIN 13843 13879
FT DOMAIN 13879 13915
FT DOMAIN 13915 13951
FT DOMAIN 13951 13987
FT DOMAIN 13987 14023
FT DOMAIN 14023 14059
FT DOMAIN 14059 14095
FT DOMAIN 14095 14131
FT DOMAIN 14131 14167
FT DOMAIN 14167 14203
FT DOMAIN 14203 14239
FT DOMAIN 14239 14275
FT DOMAIN 14275 14311
FT DOMAIN 14311 14347
FT DOMAIN 14347 14383
FT DOMAIN 14383 14419
FT DOMAIN 14419 14455
FT DOMAIN 14455 14491
FT DOMAIN 14491 14527
FT DOMAIN 14527 14563
FT DOMAIN 14563 14599
FT DOMAIN 14599 14635
FT DOMAIN 14635 14671
FT DOMAIN 14671 14707
FT DOMAIN 14707 14743
FT DOMAIN 14743 14779
FT DOMAIN 14779 14815
FT DOMAIN 14815 14851
FT DOMAIN 14851 14887
FT DOMAIN 14887 14923
FT DOMAIN 14923 14959
FT DOMAIN 14959 15000


```

FT	DOMAIN	1349	1385	EGF-LIKE 35.
FT	DOMAIN	1388	1427	EGF-LIKE 36.
FT	REPEAT	1446	1481	LIN/NOTCH 1.
FT	REPEAT	1482	1523	LIN/NOTCH 2.
FT	REPEAT	1524	1563	LIN/NOTCH 3.
FT	REPEAT	1928	1957	ANK 1.
FT	REPEAT	1961	1991	ANK 2.
FT	REPEAT	1995	2024	ANK 3.
FT	REPEAT	2028	2057	ANK 4.
FT	REPEAT	2061	2090	ANK 5.
FT	DOMAIN	1576	1579	POLY-VAL.
FT	DOMAIN	1662	1665	POLY-ARG.
FT	DOMAIN	1729	1732	POLY-PRO.
FT	DOMAIN	1741	1744	POLY-ALA.
FT	DOMAIN	1902	1905	POLY-GLU.
FT	DOMAIN	2260	2263	POLY-GLY.
FT	DOMAIN	2404	2407	POLY-GLN.
FT	DOMAIN	2411	2418	POLY-PRO.
FT	DISULFID	24	37	BY SIMILARITY.
FT	DISULFID	31	46	BY SIMILARITY.
FT	DISULFID	48	57	BY SIMILARITY.
FT	DISULFID	63	74	BY SIMILARITY.
FT	DISULFID	68	87	BY SIMILARITY.
FT	DISULFID	89	98	BY SIMILARITY.
FT	DISULFID	106	117	BY SIMILARITY.
FT	DISULFID	111	127	BY SIMILARITY.
FT	DISULFID	129	138	BY SIMILARITY.
FT	DISULFID	144	155	BY SIMILARITY.
FT	DISULFID	149	164	BY SIMILARITY.
FT	DISULFID	166	175	BY SIMILARITY.
FT	DISULFID	182	195	BY SIMILARITY.
FT	DISULFID	189	204	BY SIMILARITY.
FT	DISULFID	206	215	BY SIMILARITY.
FT	DISULFID	222	233	BY SIMILARITY.
FT	DISULFID	227	243	BY SIMILARITY.
FT	DISULFID	245	254	BY SIMILARITY.
FT	DISULFID	261	272	BY SIMILARITY.
FT	DISULFID	266	281	BY SIMILARITY.
FT	DISULFID	283	292	BY SIMILARITY.
FT	DISULFID	299	312	BY SIMILARITY.
FT	DISULFID	306	321	BY SIMILARITY.
FT	DISULFID	323	332	BY SIMILARITY.
FT	DISULFID	339	350	BY SIMILARITY.
FT	DISULFID	344	359	BY SIMILARITY.
FT	DISULFID	361	370	BY SIMILARITY.
FT	DISULFID	376	387	BY SIMILARITY.
FT	DISULFID	381	398	BY SIMILARITY.
FT	DISULFID	400	409	BY SIMILARITY.
FT	DISULFID	416	429	BY SIMILARITY.
FT	DISULFID	423	438	BY SIMILARITY.
FT	DISULFID	440	449	BY SIMILARITY.
FT	DISULFID	456	467	BY SIMILARITY.
FT	DISULFID	461	476	BY SIMILARITY.
FT	DISULFID	478	487	BY SIMILARITY.
FT	DISULFID	494	505	BY SIMILARITY.
FT	DISULFID	499	514	BY SIMILARITY.
FT	DISULFID	516	525	BY SIMILARITY.
FT	DISULFID	532	543	BY SIMILARITY.
FT	DISULFID	537	552	BY SIMILARITY.
FT	DISULFID	554	563	BY SIMILARITY.
FT	DISULFID	570	580	BY SIMILARITY.
FT	DISULFID	575	589	BY SIMILARITY.
FT	DISULFID	591	600	BY SIMILARITY.
FT	DISULFID	607	618	BY SIMILARITY.
FT	DISULFID	612	627	BY SIMILARITY.
FT	DISULFID	629	638	BY SIMILARITY.
FT	DISULFID	645	655	BY SIMILARITY.
FT	DISULFID	650	664	BY SIMILARITY.
FT	DISULFID	666	675	BY SIMILARITY.
FT	DISULFID	682	693	BY SIMILARITY.
FT	DISULFID	687	702	BY SIMILARITY.
FT	DISULFID	704	713	BY SIMILARITY.
FT	DISULFID	720	730	BY SIMILARITY.

FT	DISULFID	725	739	BY SIMILARITY.
FT	DISULFID	741	750	BY SIMILARITY.
FT	DISULFID	757	768	BY SIMILARITY.
FT	DISULFID	762	777	BY SIMILARITY.
FT	DISULFID	769	788	BY SIMILARITY.
FT	DISULFID	795	806	BY SIMILARITY.
FT	DISULFID	800	815	BY SIMILARITY.
FT	DISULFID	817	826	BY SIMILARITY.
FT	DISULFID	833	844	BY SIMILARITY.
FT	DISULFID	838	855	BY SIMILARITY.
FT	DISULFID	857	867	BY SIMILARITY.
FT	DISULFID	874	885	BY SIMILARITY.
FT	DISULFID	879	894	BY SIMILARITY.
FT	DISULFID	896	905	BY SIMILARITY.
FT	DISULFID	912	923	BY SIMILARITY.
FT	DISULFID	917	932	BY SIMILARITY.
FT	DISULFID	934	943	BY SIMILARITY.
FT	DISULFID	988	999	BY SIMILARITY.
FT	DISULFID	993	1008	BY SIMILARITY.
FT	DISULFID	1010	1019	BY SIMILARITY.
FT	DISULFID	1026	1037	BY SIMILARITY.
FT	DISULFID	1031	1046	BY SIMILARITY.
FT	DISULFID	1048	1057	BY SIMILARITY.
FT	DISULFID	1064	1075	BY SIMILARITY.
FT	DISULFID	1069	1084	BY SIMILARITY.
FT	DISULFID	1086	1095	BY SIMILARITY.
FT	DISULFID	1102	1123	BY SIMILARITY.
FT	DISULFID	1117	1132	BY SIMILARITY.
FT	DISULFID	1134	1143	BY SIMILARITY.

Query Match 2.1%; Score 10; DB 1; Length 2444;
Best Local Similarity 100.0%; Pred. No. 0.097;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps

Qy	81	NPCHNGTGTC	90
Qy	81	NPCHNGTGTC	90
Db	685	NPCHNGTGTC	694

RESULT 5
MFGM_HUMAN
ID MFGM_HUMAN STANDARD; PRT; 387 AA.
AC Q08431;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE LACTADHERIN PRECURSOR (MILK FAT GLOBULE-BSF FACTOR 8) (MFG-B8) (HMFG)
DE (BREAST EPITHELIAL ANTIGEN BA46) (MFGM) [CONTAINS: MEDIN].
GN MFG8.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Breast, and Breast carcinoma;
RX MEDLINE=96213908; PubMed=8639264;
RA Couto J.R., Taylor M.R., Godwin S.G., Ceriani R.L., Peterson J.A.;
RT "Cloning and sequence analysis of human breast epithelial antigen
RT BA46 reveals an RGD cell adhesion sequence presented on an epidermal
RT growth factor-like domain.";
RL DNA Cell Biol. 15:281-286(1996).
RN [2]
RP SEQUENCE OF 170-387 FROM N.A.
RC TISSUE=Mammary gland;
RX MEDLINE=91371351; PubMed=1909932;
RA Larocca D., Peterson J.A., Urrea R., Kuniyoshi J., Bistrain A.M.,
RA Ceriani R.L.;
RT "A Mr 46,000 human milk fat globule protein that is highly expressed
RT in human breast tumors contains factor VIII-like domains.";
RL Cancer Res. 51:4994-4998(1991).
RN [3]
RP PARTIAL SEQUENCE, AND CHARACTERIZATION.

RC TISSUE-Milk;
RX MEDLINE=98194924; PubMed=9535276;
RA Giuffrida M.G., Cavaletto M., Giunta C., Conti A.,
RA Godovac-Zimmermann J.;
RT "Isolation and characterization of full and truncated forms of human
RT breast carcinoma protein BA46 from human milk fat globule membranes.";
RL J. Protein Chem. 17:143-148(1998).
RN [4]
RP SEQUENCE OF 268-317, AND IDENTIFICATION OF MEDIN.
RX MEDLINE=99342076; PubMed=10411933;
RA Haegggqvist B., Naeslund J., Sletten K., Westermark G.T., Mucchiano G.,
RA Tjernberg L.O., Nordstedt C., Engstroem U., Westermark P.;
RT "Medin: an integral fragment of aortic smooth muscle cell-produced
RT lactadherin forms the most common human amyloid.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:8669-8674(1999).
RN [5]
RP CHARACTERIZATION.
RX MEDLINE=97405885; PubMed=9260929;
RA Taylor M.R., Couto J.R., Scallan C.D., Ceriani R.L., Peterson J.A.;
RT "Lactadherin (formerly BA46), a membrane-associated glycoprotein
RT expressed in human milk and breast carcinomas, promotes Arg-Gly-Asp
RT (RGD)-dependent cell adhesion.";
RL DNA Cell Biol. 16:861-869(1997).
CC -!- FUNCTION: MAY BE INVOLVED IN PHOSPHOLIPID BINDING. BINDS
CC SPECIFICALLY TO ROTAVIRUS AND INHIBITS ITS REPLICATION.
CC -!- FUNCTION: MEDIN IS THE MAIN CONSTITUENT OF AORTIC MEDIAL AMYLOID.
CC -!- SUBCELLULAR LOCATION: PERIPHERAL MEMBRANE PROTEIN.
CC -!- TISSUE SPECIFICITY: MAMMARY EPITHELIAL CELL SURFACES AND AORTIC
CC MEDIA. OVEREXPRESSED IN SEVERAL CARCINOMAS.
CC -!- PTM: MEDIN HAS A RAGGED N-TERMINUS WITH MINOR SPECIES STARTING AT
CC AMINO ACID 264 AND 273.
CC -!- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -!- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@lsb-sib.ch).
CC
DR EMBL; U58516; AAC50549.1; -;
DR EMBL; S56151; AAB19771.1; -;
DR MIM; 602281; -;
DR InterPro; IPR000421; -;
DR InterPro; IPR000561; -;
DR Pfam; PF00008; EGF_1;
DR Pfam; PF00754; F5_F8_type_C; 2.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01285; FA58C_1; 2.
DR PROSITE; PS01286; FA58C_2; 2.
DR KW Signal; Glycoprotein; Milk; Repeat; EGF-like domain; Amyloid.
FT SIGNAL 1 23
FT CHAIN 24 387 LACTADHERIN.
FT CHAIN 202 387 LACTADHERIN; SHORT FORM.
FT CHAIN 268 317 MEDIN.
FT DOMAIN 70 225 EGF-LIKE.
FT DOMAIN 230 387 F5/8 TYPE C 1.
FT SITE 46 38 F5/8 TYPE C 2.
FT DISULFID 27 38 CELL ATTACHMENT SITE (POTENTIAL).
FT DISULFID 32 55 BY SIMILARITY.
FT DISULFID 57 66 BY SIMILARITY.
FT DISULFID 70 225 BY SIMILARITY.
FT DISULFID 212 216 BY SIMILARITY.
FT DISULFID 230 387 BY SIMILARITY.
FT CARBOHYD 238 238 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 325 325 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 329 329 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 350 350 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT SEQUENCE 387 AA; 43123 MW; 2EE6571DEC83782D CRC64;

Query Match 1.9%; Score 9; DB 1; Length 387;
Best Local Similarity 100.0%; Pred. No. 0.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 388 VTGIITQGA 396
| | | | | | | | | |
Db 299 VTGIITQGA 307

RESULT 6
MFGM_BOVIN STANDARD; PRT; 427 AA.
ID MFGM_BOVIN Q27959; P79344;
AC Q95114; Q27959; P79344;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE LACTADHERIN PRECURSOR (MILK FAT GLOBULE-EGF FACTOR 8) (MFG-E8)
DE (MGP57/53) (PAS-6/PAS-7 GLYCOPROTEIN) (MFGM) (SPERM SURFACE PROTEIN
DE SP47) (BP47) (COMPONENTS 15/16).
GN MFGES.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A., PARTIAL SEQUENCE, AND CARBOHYDRATE-LINKAGE SITES.
RC STRAIN-HOLSTEIN; TISSUE-Mammary gland;
RX MEDLINE=97008954; PubMed=8856064;
RA Hvarregaard J., Andersen M.H., Berglund L., Rasmussen J.T.,
RA Petersen T.E.;
RT "Characterization of glycoprotein PAS-6/7 from membranes of bovine
RT milk fat globules.";
RL Eur. J. Biochem. 240:628-636(1996).
RN [2]
RP SEQUENCE OF 18-427 FROM N.A.
RC TISSUE-Mammary gland;
RX MEDLINE=96125736; PubMed=8541316;
RA Aoki N., Kishi M., Taniguchi Y., Adachi T., Nakamura R.,
RA Matsuda T.;
RT "Molecular cloning of glycoprotein antigens MGP57/53 recognized by
RT monoclonal antibodies raised against bovine milk fat globule
RT membrane.";
RL Biochim. Biophys. Acta 1245:385-391(1995).
RN [3]
RP SEQUENCE OF 19-427 FROM N.A.
RC TISSUE-Testis;
RA Ensslin M.A.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 140-146; 174-187; 233-246 AND 422-427.
RC TISSUE-Milk;
RX MEDLINE=93250576; PubMed=8485470;
RA Mather I.H., Banghart L.R., Lane W.S.;
RT "The major fat-globule membrane proteins, bovine components 15/16 and
RT guinea-pig GP 55, are homologous to MGF-E8, a murine glycoprotein
RT containing epidermal growth factor-like and factor V/VIII-like
RT sequences.";
RL Biochem. Mol. Biol. Int. 29:545-554(1993).
CC -!- FUNCTION: PROBABLY ASSOCIATES WITH PHOSPHOLIPIDS ON THE SURFACE OF
CC MAMMARY EPITHELIAL CELLS AND MILK FAT GLOBULES. ZONA PELLUCIDA-
CC BINDING PROTEIN.
CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
CC SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING. THE SHORT FORM
CC LACKS 53 AMINO ACIDS WITHIN THE F5/8 TYPE C 1 DOMAIN.
CC -!- TISSUE SPECIFICITY: MILK AND SPERMATOZOAN.
CC -!- PTM: THE 2 O-LINKED GLYCANS CONSIST OF GAL, GLCNAC AND FUC, WITH
CC PROBABLY FUC AS REDUCING TERMINAL SUGAR.
CC -!- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
CC -!- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; X91895; CAA62997.1; -;
EMBL; S80643; AAB35894.2; -;
EMBL; Y11719; CAA72406.1; -;
HSP; P00740; LIXA.
InterPro: IPR000421; -;
InterPro: IPR000561; -;
Pfam; PF00754; F5_F8_type_C; 2.
PROSITE; PS00022; EGF_1; 2.
PROSITE; PS01186; EGF_2; 2.
PROSITE; PS01285; FA58C_1; 2.
PROSITE; PS01286; FA58C_2; 2.
Signal; Glycoprotein; Milk; Repeat; EGF-like domain;
Alternative splicing;
FT SIGNAL 1 18
FT CHAIN 19 427 LACTADHERIN.
FT DOMAIN 20 59 EGF-LIKE 1.
FT DOMAIN 62 106 EGF-LIKE 2.
FT DOMAIN 109 265 F5/8 TYPE C 1.
FT DOMAIN 270 427 F5/8 TYPE C 2.
FT SITE 85 87 CELL ATTACHMENT SITE (POTENTIAL).
FT DISULFID 24 35 BY SIMILARITY.
FT DISULFID 29 47 BY SIMILARITY.
FT DISULFID 49 58 BY SIMILARITY.
FT DISULFID 66 77 BY SIMILARITY.
FT DISULFID 71 94 BY SIMILARITY.
FT DISULFID 96 105 BY SIMILARITY.
FT DISULFID 109 265
FT DISULFID 252 256
FT DISULFID 270 427
FT CARBOHYD 27 27 O-LINKED (FUC. . .) (IN PAS-6).
FT CARBOHYD 34 34 O-LINKED (FUC. . .) (IN PAS-7).
FT CARBOHYD 59 59 N-LINKED (GLCNAC. . .) (HYBRID) (IN PAS-6 AND PAS-7).
FT CARBOHYD 227 227 N-LINKED (GLCNAC. . .) (HIGH MANNOSE) (IN PAS-6).
FT VARSPLIC 169 221 MISSING (IN SHORT ISOFORM).
FT CONFLICT 19 19 A -> F (IN REF. 1).
FT CONFLICT 28 28 L -> Q (IN REF. 1).
FT SEQUENCE 427 AA; 47411 MW; 4CBEE3A1DC4EB24 CRC64;

Query Match 1.9%; Score 9; DB 1; Length 427;
Best Local Similarity 100.0%; Pred. No. 0.22;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 388 VTGIITQGA 396
| | | | | | | | | |
Db 339 VTGIITQGA 347

RESULT 7
MFGM_MOUSE STANDARD; PRT; 463 AA.
ID MFGM_MOUSE AC 21956; P97800;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE LACTADHERIN PRECURSOR (MILK FAT GLOBULE-EGF FACTOR 8) (MFG-E8) (MFGM)
DE (SPERM SURFACE PROTEIN SP47) (MP47).
GN MFGEB.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;

RN SEQUENCE FROM N.A., AND SEQUENCE OF 23-35.
RP TISSUE=Mammary gland;
RX MEDLINE=91046008; PubMed=2122462;
RA Stubbs J.D., Lekutis C., Singer K.L., Bui A., Yuzuki D.,
RA Srinivasan U., Parry G.;
RT "cDNA cloning of a mouse mammary epithelial cell surface protein
reveals the existence of epidermal growth factor-like domains linked
to factor VIII-like sequences.";
Proc. Natl. Acad. Sci. U.S.A. 87:8417-8421(1990).
[2]
RN SEQUENCE OF 23-463 FROM N.A.
RP TISSUE=Testis;
RA Ensslin M.A.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
CC !- FUNCTION: MAY BE INVOLVED IN PHOSPHOLIPID BINDING. ZONA PELLUCIDA-
CC !- BINDING PROTEIN
CC !- SUBCELLULAR LOCATION: PERIPHERAL MEMBRANE PROTEIN.
CC !- TISSUE SPECIFICITY: MAMMARY EPITHELIAL CELL SURFACES AND
CC SPERMATOZOAN.
CC !- DEVELOPMENTAL STAGE: MRNA EXPRESSION IS DETECTABLE IN MAMMARY
CC TISSUE FROM NONPREGNANT ANIMALS & MAXIMAL IN THE LACTATING GLAND.
CC !- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
CC !- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

EMBL; M38337; AAA39534.1; -;
PIR; A36479; A36479.
HSP; P00740; LIXA.
MGD; MGI:102768; Mfge8.
InterPro: IPR000421; -;
InterPro: IPR000561; -;
InterPro: IPR001438; -;
Pfam; PF00008; EGF; 2.
Pfam; PF00754; F5_F8_type_C; 2.
PRINTS; PR00010; EGFBL00D.
PROSITE; PS00022; EGF_1; 2.
PROSITE; PS01186; EGF_2; 2.
PROSITE; PS01285; FA58C_1; 2.
PROSITE; PS01286; FA58C_2; 2.
Signal; Glycoprotein; Repeat; EGF-like domain; Milk.
FT SIGNAL 1 22
FT CHAIN 23 463 LACTADHERIN.
FT DOMAIN 24 61 EGF-LIKE 1.
FT DOMAIN 64 108 EGF-LIKE 2.
FT DOMAIN 148 303 F5/8 TYPE C 1.
FT DOMAIN 308 463 F5/8 TYPE C 2.
FT SITE 87 89 CELL ATTACHMENT SITE (POTENTIAL).
FT DISULFID 28 39 BY SIMILARITY.
FT DISULFID 33 49 BY SIMILARITY.
FT DISULFID 51 60 BY SIMILARITY.
FT DISULFID 68 79 BY SIMILARITY.
FT DISULFID 73 96 BY SIMILARITY.
FT DISULFID 98 107 BY SIMILARITY.
FT DISULFID 148 303 BY SIMILARITY.
FT DISULFID 290 294 BY SIMILARITY.
FT DISULFID 308 463 BY SIMILARITY.
FT CARBOHYD 61 61 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 266 266 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 316 316 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 426 426 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 30 30 S -> F (IN REF. 2).
FT CONFLICT 35 35 N -> D (IN REF. 1; AA SEQUENCE).
FT CONFLICT 110 147 EFTNYNLDGEYMTFAVPTAVPTAPDPDLSNNLASR ->
G (IN REF. 2).

FT CONFLICT 168 168 Y -> S (IN REF. 2).
FT CONFLICT 196 196 H -> T (IN REF. 2).
FT CONFLICT 309 309 L -> S (IN REF. 2).
FT CONFLICT 395 395 E -> A (IN REF. 2).
SQ SEQUENCE 463 AA; 51465 MW; D78B6C6EFBBA724D CRC64;

Query Match 1.9%; Score 9; DB 1; Length 463;
Best Local Similarity 100.0%; Pred. No. 0.24;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 388 VTGIITQGA 396
|||||||
Db 377 VTGIITQGA 385

RESULT 8
NRPL_RAT STANDARD; PRT; 922 AA.
AC Q9QWJ9;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE NEUROFILIN-1 PRECURSOR (VASCULAR ENDOTHELIAL CELL GROWTH FACTOR 165 RECEPTOR).
DE RECEPTOR.
GN NRPL.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SPRAGUE-DAWLEY;
RX MEDLINE=97433085; PubMed=9288754;
RA Kolodkin A.L., Levengood D.V., Rowe E.G., Tai Y.-T., Giger R.J.,
RA Ginty D.D.;
RT "Neuropilin is a semaphorin III receptor.";
RL Cell 90:753-762(1997).

CC -1- FUNCTION: RECEPTOR INVOLVED IN THE DEVELOPMENT OF THE
CARDIOVASCULAR SYSTEM, IN ANGIOGENESIS, IN THE FORMATION OF
CERTAIN NEURONAL CIRCUITS AND IN ORGANOGENESIS OUTSIDE THE NERVOUS
SYSTEM. IT MEDIATES THE CHEMOREPULSANT ACTIVITY OF SEMAPHORINS. IT
BINDS TO SEMAPHORIN 3A, THE PLGF-2 ISOFORM OF PGF, THE VEGF-165
ISOFORM OF VEGF AND VEGF-B. COEXPRESSION WITH KDR RESULTS IN
INCREASED VEGF-165 BINDING TO KDR AS WELL AS INCREASED CHEMOTAXIS.
IT MAY REGULATE VEGF-INDUCED ANGIOGENESIS (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- TISSUE SPECIFICITY: FOUND IN THE EMBRYONIC NERVOUS SYSTEM.
CC -1- SIMILARITY: BELONGS TO THE NEUROFILIN FAMILY.
CC -1- SIMILARITY: CONTAINS 2 CUB DOMAINS.
CC -1- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.
CC -1- SIMILARITY: CONTAINS 1 MAM DOMAIN.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

DR EMBL; AF016296; AAC53337.1;
DR InterPro; IPR000421;
DR InterPro; IPR000859;
DR InterPro; IPR000998;
DR Pfam; PF00431; CUB; 2.
DR Pfam; PF00629; MAM; 1.
DR Pfam; PF00754; F5_f8_type_C; 2.
DR PRINTS; PR00020; MAMDOMAIN.
DR PROSITE; PS01180; CUB; 2.
DR PROSITE; PS01285; FA58C.1; 2.
DR PROSITE; PS01286; FA58C.2; 2.
DR PROSITE; PS00740; MAM_1; 1.

DR PROSITE; PS50060; MAM_2; 1.
KW Transmembrane; Glycoprotein; Neurone; Signal; Repeat; Receptor.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 922 NEUROFILIN-1.
FT DOMAIN 22 855 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 856 880 POTENTIAL.
FT DOMAIN 881 922 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 27 141 CUB 1.
FT DOMAIN 147 265 CUB 2.
FT DOMAIN 275 424 F5/8 TYPE C 1.
FT DOMAIN 431 583 F5/8 TYPE C 2.
FT DOMAIN 645 811 MAM.
FT DISULFID 27 54 PROBABLE.
FT DISULFID 82 104 PROBABLE.
FT DISULFID 147 173 PROBABLE.
FT DISULFID 206 228 PROBABLE.
FT DISULFID 275 424 BY SIMILARITY.
FT DISULFID 431 583 BY SIMILARITY.
FT CARBOHYD 150 150 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 261 261 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 300 300 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 522 522 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 841 841 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 922 AA; 103082 MW; CC6F82AD098B0F2E CRC64;

Query Match 1.9%; Score 9; DB 1; Length 922;
Best Local Similarity 100.0%; Pred. No. 0.45;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 307 LRMEILGCE 315
|||||||
Db 576 LRMEILGCE 584

RESULT 9
NRPL_HUMAN STANDARD; PRT; 923 AA.
ID NRPL_HUMAN
AC O14786; O60461;
DT 30-MAY-2000 (Rel. 39, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE NEUROFILIN-1 PRECURSOR (VASCULAR ENDOTHELIAL CELL GROWTH FACTOR 165 RECEPTOR).
DE RECEPTOR.
GN NRPL OR NRP OR VEGF165R.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (MEMBRANE-BOUND ISOFORM).
RX MEDLINE=97433084; PubMed=9288753;
RA He Z., Tessier-Lavigne M.;
RT "Neuropilin is a receptor for the axonal chemorepellent semaphorin III.";
RL Cell 90:739-751(1997).
RN [2]
RP SEQUENCE FROM N.A. (MEMBRANE-BOUND ISOFORM), AND SEQUENCE OF 22-39.
RC TISSUE-Breast;
RX MEDLINE=98188099; PubMed=9529250;
RA Soker S., Takashima S., Miao H.-Q., Neufeld G., Klagsbrun M.;
RT "Neuropilin-1 is expressed by endothelial and tumor cells as an isoform-specific receptor for vascular endothelial growth factor.";
RL Cell 92:735-745(1998).
RN [3]
RP SEQUENCE FROM N.A. (SOLUBLE/SNRP1 ISOFORM), AND SEQUENCE OF 22-31.
RC TISSUE-Prostatic adenocarcinoma;
RX MEDLINE=20183929; PubMed=10688880;
RA Gagnon M.L., Bielenberg D.R., Gechtman Z., Miao H.-Q., Takashima S.,
RA Soker S., Klagsbrun M.;
RT "Identification of a natural soluble neuropilin-1 that binds vascular endothelial growth factor: In vivo expression and antitumor activity.";
RT

CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>)
 CC or send an email to license@isb-sib.ch)

```

CC EMBL; D50086; BAA08789.1; -.
CC MGD; MGI:106206; Nrp.
DR InterPro; IPR000421; -.
DR InterPro; IPR000859; -.
DR InterPro; IPR000998; -.
DR Pfam; PF00431; CUB; 2.
DR Pfam; PF00754; F5_F8_type_C; 2.
DR Pfam; PF00629; MAM; 1.
DR PROSITE; PS01180; CUB; 2.
DR PROSITE; PS01285; FAS8C.1; 2.
DR PROSITE; PS01286; FAS8C.2; 2.
DR PROSITE; PS00740; MAM.1; 1.
DR PROSITE; PS00060; MAM.2; 1.
KW Transmembrane; Glycoprotein; Neutrone; Signal; Repeat; Receptor.
FT SIGNAL 1 21
FT CHAIN 22 923
FT DOMAIN 22 856
FT TRANSMEM 857 879
FT DOMAIN 880 923
FT DOMAIN 27 141
FT DOMAIN 147 265
FT DOMAIN 275 424
FT DOMAIN 431 583
FT DOMAIN 645 811
FT DISULFID 27 54
FT DISULFID 82 104
FT DISULFID 147 173
FT DISULFID 206 228
FT DISULFID 275 424
FT DISULFID 431 583
FT CARBOHYD 150 150
FT CARBOHYD 261 261
FT CARBOHYD 300 300
FT CARBOHYD 522 522
FT CARBOHYD 842 842
SQ SEQUENCE 923 AA; 103020 MW; 0644B8A170796808 CRC64;

Query Match 1.9%; Score 9; DB 1; Length 923;
Best Local Similarity 100.0%; Pred. No. 0.45;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 307 LRMEILGCE 315
Db 576 LRMEILGCE 584

RESULT 11
SERR_DROME
ID SERR_DROME STANDARD; PRT; 1408 AA.
AC P18168;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE SERATE PROTEIN PRECURSOR (BEADED PROTEIN).
GN SER OR BD.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=91347903; PubMed=1840519;
RA Thomas U., Speicher S.A., Knust E.;
RT "The Drosophila gene Serrate encodes an EGF-like transmembrane
RT protein with a complex expression pattern in embryos and wing
RT discs.";
RL Development 111:749-761(1991).

```

```

RN RP SEQUENCE FROM N.A.
RX MEDLINE=91099666; PubMed=2125287;
RA Fleming R.J., Scottgale T.N., Diederich R.J., Artavanis-Tsakonas S.;
RT "The gene Serrate encodes a putative EGF-like transmembrane protein
RT essential for proper ectodermal development in Drosophila
RT melanogaster."
RL Genes Dev. 4:2188-2201(1990).
CC -1- FUNCTION: ESSENTIAL FOR PROPER ECTODERMAL DEVELOPMENT. SERRATE
CC MAY REPRESENT AN ELEMENT IN A NETWORK OF INTERACTING MOLECULES
CC OPERATING AT THE CELL SURFACE DURING THE DIFFERENTIATION OF
CC CERTAIN TISSUES.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- TISSUE SPECIFICITY: APPEARS TO BE RESTRICTED EXCLUSIVELY TO
CC CELLS OF ECTODERMAL ORIGIN.
CC -1- MISCELLANEOUS: SEPARATION OF NEUROBLASTS FROM THE ECTODERM INTO
CC THE INNER PART OF EMBRYO IS ONE OF THE FIRST STEPS OF CNS
CC DEVELOPMENT IN INSECTS, THIS PROCESS IS UNDER CONTROL OF THE
CC NEUROGENIC GENES.
CC -1- MISCELLANEOUS: NOTCH AND SERRATE MAY INTERACT AT THE PROTEIN
CC LEVEL, IT IS CONCEIVABLE THAT THE SERRATE AND DELTA PROTEINS MAY
CC COMPETE FOR BINDING WITH THE NOTCH PROTEIN.
CC -1- SIMILARITY: CONTAINS 14 EGF-LIKE DOMAINS.
CC -1- SIMILARITY: BELONGS TO THE DELTA/SERRATE/JAGGED FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X56811; CAA40148.1; -.
DR EMBL; M35759; AAA28938.1; -.
DR PIR; A36666; A36666.
DR PIR; S16878; S16878.
DR HSP; P00743; LWHE.
DR FlyBase; FBgn0004197; Ser.
DR InterPro; IPR000152; -.
DR InterPro; IPR000561; -.
DR InterPro; IPR001438; -.
DR InterPro; IPR001774; -.
DR InterPro; IPR001881; -.
DR Pfam; PF01414; DSL; 1.
DR Pfam; PF00008; EGF; 11.
DR PRINTS; PR00010; EGFBL00.
DR PROSITE; PS00010; ASX_HYDROXYL; 7.
DR PROSITE; PS00022; EGF_1; 14.
DR PROSITE; PS01186; EGF_2; 8.
DR PROSITE; PS01187; EGF_CA; 5.
KW Differentiation; Repeat; EGF-like domain; Transmembrane;
KW Glycoprotein; Signal.
FT SIGNAL 1 83
FT CHAIN 84 1408
FT DOMAIN 84 1223
FT TRANSMEM 1224 1249
FT DOMAIN 1250 1408
FT DOMAIN 284 317
FT DOMAIN 315 349
FT DOMAIN 351 389
FT DOMAIN 391 489
FT DOMAIN 407 476
FT DOMAIN 491 527
FT DOMAIN 529 609
FT DOMAIN 611 646
FT DOMAIN 648 684
FT DOMAIN 686 721
FT DOMAIN 723 797
FT DOMAIN 737 769
FT DOMAIN 799 835
FT DOMAIN 837 877
FT DOMAIN 879 915

```

```
FT DOMAIN 917 953 EGF-LIKE 14, CALCIUM-BINDING (POTENTIAL).
FT DISULFID 288 299 BY SIMILARITY.
FT DISULFID 292 305 BY SIMILARITY.
FT DISULFID 307 316 BY SIMILARITY.
FT DISULFID 319 330 BY SIMILARITY.
FT DISULFID 325 337 BY SIMILARITY.
FT DISULFID 339 348 BY SIMILARITY.
FT DISULFID 355 367 BY SIMILARITY.
FT DISULFID 361 377 BY SIMILARITY.
FT DISULFID 379 388 BY SIMILARITY.
FT DISULFID 395 406 BY SIMILARITY.
FT DISULFID 400 477 BY SIMILARITY.
FT DISULFID 479 488 BY SIMILARITY.
FT DISULFID 495 506 BY SIMILARITY.
FT DISULFID 500 515 BY SIMILARITY.
FT DISULFID 517 526 BY SIMILARITY.
FT DISULFID 533 588 BY SIMILARITY.
FT DISULFID 582 597 BY SIMILARITY.
FT DISULFID 599 608 BY SIMILARITY.
FT DISULFID 615 625 BY SIMILARITY.
FT DISULFID 619 634 BY SIMILARITY.
FT DISULFID 636 645 BY SIMILARITY.
FT DISULFID 652 663 BY SIMILARITY.
FT DISULFID 657 672 BY SIMILARITY.
FT DISULFID 674 683 BY SIMILARITY.
FT DISULFID 690 700 BY SIMILARITY.
FT DISULFID 695 709 BY SIMILARITY.
FT DISULFID 711 720 BY SIMILARITY.
FT DISULFID 803 814 BY SIMILARITY.
FT DISULFID 808 823 BY SIMILARITY.
FT DISULFID 825 834 BY SIMILARITY.
FT DISULFID 841 852 BY SIMILARITY.
FT DISULFID 846 865 BY SIMILARITY.
FT DISULFID 867 876 BY SIMILARITY.
FT DISULFID 883 894 BY SIMILARITY.
FT DISULFID 888 903 BY SIMILARITY.
FT DISULFID 905 914 BY SIMILARITY.
FT DISULFID 921 932 BY SIMILARITY.
FT DISULFID 926 941 BY SIMILARITY.
FT DISULFID 943 952 BY SIMILARITY.
FT CARBOHYD 152 196 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 196 247 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 247 331 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 331 331 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 412 412 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 452 452 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 558 558 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 739 739 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 965 965 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 977 977 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1004 1004 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1030 1030 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1150 1150 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 14 17 MISSING (IN REF. 2).
FT CONFLICT 27 27 P > A (IN REF. 2).
FT CONFLICT 1352 1352 T -> S (IN REF. 2).
SQ SEQUENCE 1408 AA; 150660 MW; 569DA4270A9C7840 CRC64;

Query Match 1.9%; Score 9; DB 1; Length 1408;
Best Local Similarity 100.0%; Pred. No. 0.65;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 82 PCHNGGTCE 90
Db 887 PCHNGGTCE 895
|||||||
|||||||

RESULT 12
CRB_DRCWE
ID CRB_DROME STANDARD; PRT; 2139 AA.
AC P10040;
DT 01-MAR-1989 (Rel. 10, Created)
```

```
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE CRUMBS PROTEIN PRECURSOR (95F).
GN CRB
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OREGON-R; TISSUE=Embryo;
RX MEDLINE=90263104; PubMed=2344615;
RA Tepass U., Theres C., Knust E.;
RT "Crumbs encodes an EGF-like protein expressed on apical membranes of
RT Drosophila epithelial cells and required for organization of
RT epithelia".
RL Cell 61:787-799(1990).
RN [2]
RP SEQUENCE OF 1663-1955 FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=87218537; PubMed=3107986;
RA Knust E., Dietrich U., Tepass U., Bremer K.A., Weigel D.,
RA Vaessin H., Campos-Ortega J.A.;
RT "EGF homologous sequences encoded in the genome of Drosophila
RT melanogaster, and their relation to neurogenic genes.";
RL EMBO J. 6:761-766(1987).
CC -|- FUNCTION: MAY PLAY A ROLE IN THE DEVELOPMENT OF EPITHELIA,
CC POSSIBLY FOR THE ESTABLISHMENT AND/OR MAINTENANCE OF CELL
CC POLARITY. IT MAY ACT AS A SIGNAL.
CC -|- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -|- PTM: PHOSPHORYLATED IN THE CYTOPLASMIC DOMAIN (POTENTIAL).
CC -|- SIMILARITY: CONTAINS 29 EGF-LIKE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M33753; AA28428.1; ALT_SEQ.
CC EMBL; X05144; CAA28793.1; -.
CC PIR; B26637; B26637.
CC PIR; A35672; A35672.
CC HSP; P00740; 11XA.
CC FlyBase; FBgn000368; crb.
CC InterPro; IPR000152; -.
CC InterPro; IPR000561; -.
CC InterPro; IPR001438; -.
CC InterPro; IPR001791; -.
CC InterPro; IPR001881; -.
CC Pfam; PF00008; EGF; 27.
CC Pfam; PF00054; laminin_G; 3.
CC PRINTS; PR00010; EGFBLD.
CC PROSITE; PS00010; ASX_HYDROXYL; 15.
CC PROSITE; PS00022; EGF_1; 26.
CC PROSITE; PS01186; EGF_2; 17.
CC PROSITE; PS01187; EGF_CA; 15.
KW Differentiation; Repeat; EGF-like domain; Transmembrane;
KW Glycoprotein; Signal; Phosphorylation.
FT SIGNAL 1 90
FT CHAIN 91 2139 CRUMBS PROTEIN.
FT DOMAIN 91 2084 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 2085 2111 POTENTIAL.
FT DOMAIN 2112 2139 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 267 303 EGF-LIKE 1.
FT DOMAIN 306 343 EGF-LIKE 2.
FT DOMAIN 348 386 EGF-LIKE 3.
FT DOMAIN 388 425 EGF-LIKE 4.
FT DOMAIN 427 463 EGF-LIKE 5.
FT DOMAIN 464 500 EGF-LIKE 6.
```

FT	DOMAIN	501	EGF-LIKE 7.	FT	DISULFID	930	939	BY SIMILARITY.
FT	DOMAIN	545	EGF-LIKE 8.	FT	DISULFID	946	957	BY SIMILARITY.
FT	DOMAIN	582	EGF-LIKE 9.	FT	DISULFID	952	966	BY SIMILARITY.
FT	DOMAIN	609	EGF-LIKE 10.	FT	DISULFID	968	977	BY SIMILARITY.
FT	DOMAIN	648	EGF-LIKE 11.	FT	DISULFID	984	995	BY SIMILARITY.
FT	DOMAIN	687	EGF-LIKE 12.	FT	DISULFID	989	1009	BY SIMILARITY.
FT	DOMAIN	725	EGF-LIKE 13.	FT	DISULFID	1011	1020	BY SIMILARITY.
FT	DOMAIN	763	EGF-LIKE 14.	FT	DISULFID	1211	1222	BY SIMILARITY.
FT	DOMAIN	802	EGF-LIKE 15.	FT	DISULFID	1216	1231	BY SIMILARITY.
FT	DOMAIN	840	EGF-LIKE 16.	FT	DISULFID	1233	1242	BY SIMILARITY.
FT	DOMAIN	904	EGF-LIKE 17.	FT	DISULFID	1485	1496	BY SIMILARITY.
FT	DOMAIN	942	EGF-LIKE 18.	FT	DISULFID	1490	1505	BY SIMILARITY.
FT	DOMAIN	980	EGF-LIKE 19.	FT	DISULFID	1507	1516	BY SIMILARITY.
FT	DOMAIN	1207	EGF-LIKE 20.	FT	DISULFID	1763	1774	BY SIMILARITY.
FT	DOMAIN	1243	EGF-LIKE 21.	FT	DISULFID	1768	1783	BY SIMILARITY.
FT	DOMAIN	1517	EGF-LIKE 22.	FT	DISULFID	1785	1794	BY SIMILARITY.
FT	DOMAIN	1759	EGF-LIKE 23.	FT	DISULFID	1801	1812	BY SIMILARITY.
FT	DOMAIN	1797	EGF-LIKE 24.	FT	DISULFID	1806	1821	BY SIMILARITY.
FT	DOMAIN	1835	EGF-LIKE 25.	FT	DISULFID	1823	1832	BY SIMILARITY.
FT	DOMAIN	1874	EGF-LIKE 26.	FT	DISULFID	1839	1850	BY SIMILARITY.
FT	DOMAIN	1915	EGF-LIKE 27.	FT	DISULFID	1844	1859	BY SIMILARITY.
FT	DOMAIN	1953	EGF-LIKE 28.	FT	DISULFID	1861	1870	BY SIMILARITY.
FT	DOMAIN	1991	EGF-LIKE 29.	FT	DISULFID	1878	1889	BY SIMILARITY.
FT	DOMAIN	2030	BY SIMILARITY.	FT	DISULFID	1883	1903	BY SIMILARITY.
FT	DISULFID	271	BY SIMILARITY.	FT	DISULFID	1905	1914	BY SIMILARITY.
FT	DISULFID	276	BY SIMILARITY.	FT	DISULFID	1919	1930	BY SIMILARITY.
FT	DISULFID	291	BY SIMILARITY.	FT	DISULFID	1924	1939	BY SIMILARITY.
FT	DISULFID	293	BY SIMILARITY.	FT	DISULFID	1941	1950	BY SIMILARITY.
FT	DISULFID	310	BY SIMILARITY.	FT	DISULFID	1957	1968	BY SIMILARITY.
FT	DISULFID	315	BY SIMILARITY.	FT	DISULFID	1962	1977	BY SIMILARITY.
FT	DISULFID	331	BY SIMILARITY.	FT	DISULFID	1979	1988	BY SIMILARITY.
FT	DISULFID	333	BY SIMILARITY.	FT	DISULFID	1995	2008	BY SIMILARITY.
FT	DISULFID	352	BY SIMILARITY.	FT	DISULFID	2002	2017	BY SIMILARITY.
FT	DISULFID	357	BY SIMILARITY.	FT	DISULFID	2019	2028	BY SIMILARITY.
FT	DISULFID	376	BY SIMILARITY.	FT	CARBOHYD	37	37	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	DISULFID	392	BY SIMILARITY.	FT	CARBOHYD	96	96	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	DISULFID	397	BY SIMILARITY.	FT	CARBOHYD	198	198	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	DISULFID	412	BY SIMILARITY.	FT	CARBOHYD	238	238	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	DISULFID	414	BY SIMILARITY.	FT	CARBOHYD	239	239	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	DISULFID	431	BY SIMILARITY.	FT	CARBOHYD	336	336	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	DISULFID	442	BY SIMILARITY.	FT	CARBOHYD	400	400	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	DISULFID	442	BY SIMILARITY.	FT	CARBOHYD	550	550	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	DISULFID	451	BY SIMILARITY.	FT	CARBOHYD	565	565	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	DISULFID	453	BY SIMILARITY.	FT	CARBOHYD	736	736	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	DISULFID	468	BY SIMILARITY.	FT	CARBOHYD	746	746	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	DISULFID	473	BY SIMILARITY.	FT	CARBOHYD	860	860	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	DISULFID	488	BY SIMILARITY.	FT	CARBOHYD	884	884	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	DISULFID	490	BY SIMILARITY.	FT	CARBOHYD	976	976	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	DISULFID	505	BY SIMILARITY.	FT	CARBOHYD	1102	1102	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	DISULFID	509	BY SIMILARITY.	FT	CARBOHYD	1114	1114	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	DISULFID	522	BY SIMILARITY.	FT	CARBOHYD	1138	1138	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	DISULFID	522	BY SIMILARITY.	FT	CARBOHYD			
FT	DISULFID	549	BY SIMILARITY.	FT	CARBOHYD			
FT	DISULFID	556	BY SIMILARITY.	FT	CARBOHYD			
FT	DISULFID	569	BY SIMILARITY.	FT	CARBOHYD			
FT	DISULFID	571	BY SIMILARITY.	FT	CARBOHYD			
FT	DISULFID	580	BY SIMILARITY.	FT	CARBOHYD			
FT	DISULFID	586	BY SIMILARITY.	FT	CARBOHYD			
FT	DISULFID	591	BY SIMILARITY.	FT	CARBOHYD			
FT	DISULFID	604	BY SIMILARITY.	FT	CARBOHYD			
FT	DISULFID	613	BY SIMILARITY.	FT	CARBOHYD			
FT	DISULFID	618	BY SIMILARITY.	FT	CARBOHYD			
FT	DISULFID	636	BY SIMILARITY.	FT	CARBOHYD			
FT	DISULFID	645	BY SIMILARITY.	FT	CARBOHYD			
FT	DISULFID	652	BY SIMILARITY.	FT	CARBOHYD			
FT	DISULFID	659	BY SIMILARITY.	FT	CARBOHYD			
FT	DISULFID	673	BY SIMILARITY.	FT	CARBOHYD			
FT	DISULFID	675	BY SIMILARITY.	FT	CARBOHYD			
FT	DISULFID	691	BY SIMILARITY.	FT	CARBOHYD			
FT	DISULFID	696	BY SIMILARITY.	FT	CARBOHYD			
FT	DISULFID	711	BY SIMILARITY.	FT	CARBOHYD			
FT	DISULFID	713	BY SIMILARITY.	FT	CARBOHYD			
FT	DISULFID	729	BY SIMILARITY.	FT	CARBOHYD			
FT	DISULFID	734	BY SIMILARITY.	FT	CARBOHYD			
FT	DISULFID	749	BY SIMILARITY.	FT	CARBOHYD			
FT	DISULFID	751	BY SIMILARITY.	FT	CARBOHYD			
FT	DISULFID	767	BY SIMILARITY.	FT	CARBOHYD			
FT	DISULFID	772	BY SIMILARITY.	FT	CARBOHYD			
FT	DISULFID	789	BY SIMILARITY.	FT	CARBOHYD			
FT	DISULFID	806	BY SIMILARITY.	FT	CARBOHYD			
FT	DISULFID	811	BY SIMILARITY.	FT	CARBOHYD			
FT	DISULFID	828	BY SIMILARITY.	FT	CARBOHYD			
FT	DISULFID	844	BY SIMILARITY.	FT	CARBOHYD			
FT	DISULFID	849	BY SIMILARITY.	FT	CARBOHYD			
FT	DISULFID	892	BY SIMILARITY.	FT	CARBOHYD			
FT	DISULFID	901	BY SIMILARITY.	FT	CARBOHYD			
FT	DISULFID	908	BY SIMILARITY.	FT	CARBOHYD			
FT	DISULFID	919	BY SIMILARITY.	FT	CARBOHYD			
FT	DISULFID	933	BY SIMILARITY.	FT	CARBOHYD			
FT	DISULFID	928	BY SIMILARITY.	FT	CARBOHYD			

Query Match 1.9%; Score 9; DB 1; Length 2139;
Best Local Similarity 100.0%; Pred. No. 0.95;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 48 CECPEGFAG 56
Db 2058 CECPEGFAG 2066

RESULT 13
FAB_MOUSE STANDARD; PRT: 2319 AA.
AC Q06194;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE . COAGULATION FACTOR VII PRECURSOR (PROCOAGULANT COMPONENT).
GN CF8 OR F8C.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6 X CBA; TISSUE=Liver;
 RA MEDLINE=93300511; PubMed=8314577;
 RX Elder B., Lakich D., Gitschler J.;
 RT "Sequence of the murine factor VIII CDNA.";
 RL Genomics 16:374-379(1993).
 CC -!- FUNCTION: FACTOR VIII, ALONG WITH CALCIUM AND PHOSPHOLIPID, ACTS
 CC AS A COFACTOR FOR FACTOR IXA WHEN IT CONVERTS FACTOR X TO THE
 CC ACTIVATED FORM, FACTOR XA.
 CC -!- SUBCELLULAR LOCATION: EXTRACELLULAR.
 CC -!- TISSUE SPECIFICITY: FOUND IN MOST TISSUES.
 CC -!- SIMILARITY: CONTAINS 3 F5/8 TYPE A DOMAINS; EACH IS COMPOSED OF
 CC 2 PLASTOCYANIN-LIKE REPEATS.
 CC -!- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.
 CC -!- SIMILARITY: STRONG, TO COAGULATION FACTOR V.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; L05573; AAA37385.1; -.
 CC PIR; A47004; A47004.
 CC HSP; P00451; 1CFG.
 CC MGD; MG1:88383; F8.
 CC InterPro; IPR001117; -.
 CC InterPro; IPR000421; -.
 CC Pfam; PF00754; Cu-oxidase; 3.
 CC DR PFam; PF00754; F5_F8_type.C; 2.
 CC DR PROSITE; PS00079; MULTICOPPER_OXIDASE1; 3.
 CC DR PROSITE; PS01285; FA58C_1; 2.
 CC DR PROSITE; PS01286; FA58C_2; 2.
 CC KW Blood coagulation; Repeat; Plasma; Acute phase; Calcium;
 CC Signal; Glycoprotein; Sulfatation.
 CC FT SIGNAL 1 19 POTENTIAL.
 CC FT CHAIN 20 2319 COAGULATION FACTOR VIII.
 CC FT DOMAIN 20 349 F5/8 TYPE A 1.
 CC FT DOMAIN 20 199 PLASTOCYANIN-LIKE 1.
 CC FT DOMAIN 207 349 PLASTOCYANIN-LIKE 2.
 CC FT DOMAIN 399 730 F5/8 TYPE A 2.
 CC FT DOMAIN 399 573 PLASTOCYANIN-LIKE 3.
 CC FT DOMAIN 583 730 PLASTOCYANIN-LIKE 4.
 CC FT DOMAIN 760 1640 B.
 CC FT DOMAIN 1683 2008 F5/8 TYPE A 3.
 CC FT DOMAIN 1683 1845 PLASTOCYANIN-LIKE 5.
 CC FT DOMAIN 1855 2008 PLASTOCYANIN-LIKE 6.
 CC FT DOMAIN 2008 2156 F5/8 TYPE C 1.
 CC FT DOMAIN 2161 2313 F5/8 TYPE C 2.
 CC FT SITE 391 392 CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
 CC FT SITE 759 760 CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
 CC FT SITE 1678 1679 CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
 CC FT SITE 1324 1325 CLEAVAGE (ACTIVATION) (BY SIMILARITY).
 CC FT SITE 1640 1641 CLEAVAGE (ACTIVATION) (BY SIMILARITY).
 CC FT MOD_RES 367 367 SULFATATION (BY SIMILARITY).
 CC FT MOD_RES 737 737 SULFATATION (BY SIMILARITY).
 CC FT MOD_RES 738 738 SULFATATION (BY SIMILARITY).
 CC FT MOD_RES 742 742 SULFATATION (BY SIMILARITY).
 CC FT MOD_RES 1669 1669 SULFATATION (REQUIRED FOR VWF BINDING)
 CC (BY SIMILARITY).
 CC FT MOD_RES 1687 1687 SULFATATION (BY SIMILARITY).
 CC FT DISULFID 173 199 PROBABLE.
 CC FT DISULFID 547 573 PROBABLE.
 CC FT DISULFID 1819 1845 BY SIMILARITY.
 CC FT DISULFID 2008 2156 BY SIMILARITY.
 CC FT DISULFID 2161 2313 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 61 61 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 233 233 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 259 259 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC -----

FT CARBOHYD 423 423 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 601 601 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 880 880 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 958 958 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1015 1015 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1022 1022 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1026 1026 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1044 1044 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1076 1076 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1087 1087 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1136 1136 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1161 1161 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1192 1192 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1255 1255 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1268 1268 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1273 1273 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1274 1274 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1302 1302 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1316 1316 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1340 1340 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1378 1378 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1797 1797 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 2105 2105 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 2319 AA; 266148 MW; FD054DE051DB2A01 CRC64;
 Query Match 1.9%; Score 9; DB 1; Length 2319;
 Best Local Similarity 100.0%; Pred. No. 1;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 387 KVTGIITQG 395
 DB 2226 KVTGIITQG 2234
 RESULT 14
 NOTC_BRARE
 ID NOTC_BRARE STANDARD; PRT; 2437 AA.
 AC P46530;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 GN NEUROGENIC LOCUS NOTCH HOMOLOG PROTEIN PRECURSOR.
 OS Brachydanio rerio (Zebrafish) (Zebra danio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
 OC Cypriniformes; Cyprinidae; Rasbora; Danio.
 OX NCBI_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RX MEDLINE=94128602; PubMed=8297791;
 RA Bierkamp C., Campos-Ortega J.A.;
 RT "A zebrafish homologue of the Drosophila neurogenic gene Notch and
 its pattern of transcription during early embryogenesis.";
 RL Mech. Dev. 43:87-100(1993).
 CC -!- FUNCTION: IMPLICATED IN CELL FATE SPECIFICATIONS DURING
 CC EMBRYO DEVELOPMENT. MAY BE INVOLVED IN THE FORMATION OF THE
 CC NEURAL PLATE, NOTOCHORD AND BRAIN VESICLES.
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -!- DEVELOPMENTAL STAGE: EXPRESSED IN ALL CELLS IN PREGASTRULATION
 CC STAGES. DURING GASTRULATION IS DIFFERENTIALLY EXPRESSED.
 CC ACCUMULATING PREDOMINANTLY IN THE PRECHORDAL MESODERM AND
 CC NOTOCHORD. AT THE END OF GASTRULATION, EXPRESSED ALONG THE
 CC ANTERIOR-POSTERIOR AXIS INCLUDING THE DEVELOPING NEURAL PLATE
 CC AND DIFFERENTIATING MESODERM. ALSO PRESENT IN THE DEVELOPING
 CC BRAIN AND HEAD REGIONS.
 CC -!- SIMILARITY: HIGH, WITH OTHER NOTCH-TYPE PROTEINS.
 CC -!- SIMILARITY: CONTAINS 36 EGF-LIKE DOMAINS.
 CC -!- SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
 CC -!- SIMILARITY: CONTAINS 6 ANK REPEATS.
 CC -----


```

FT DISULFID 836 853 BY SIMILARITY.
FT DISULFID 855 864 BY SIMILARITY.
FT DISULFID 871 882 BY SIMILARITY.
FT DISULFID 876 891 BY SIMILARITY.
FT DISULFID 893 902 BY SIMILARITY.
FT DISULFID 909 920 BY SIMILARITY.
FT DISULFID 914 929 BY SIMILARITY.
FT DISULFID 931 940 BY SIMILARITY.
FT DISULFID 947 958 BY SIMILARITY.
FT DISULFID 952 967 BY SIMILARITY.
FT DISULFID 969 978 BY SIMILARITY.
FT DISULFID 1023 1034 BY SIMILARITY.
FT DISULFID 1028 1043 BY SIMILARITY.
FT DISULFID 1045 1054 BY SIMILARITY.
FT DISULFID 1061 1072 BY SIMILARITY.
FT DISULFID 1066 1081 BY SIMILARITY.
FT DISULFID 1083 1092 BY SIMILARITY.
FT DISULFID 1099 1120 BY SIMILARITY.

Query Match 1.9%; Score 9; DB 1; Length 2437;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 81 NPCHNGTC 89
Db 683 NPCHNGTC 691

RESULT 15
NTCL_MOUSE STANDARD; PRT: 2531 AA.
AC Q01705;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DE 01-OCT-2000 (Rel. 40, Last annotation update)
DE NEUROGENIC LOCUS NOTCH HOMOLOG PROTEIN 1 PRECURSOR (NOTCH PROTEIN).
GN NOTCH1 OR NOTCH.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA MEDLINE=93194170; PubMed=8449489;
RA Franco del Amo F., Gendron-Maguire M., Swiatek P.J., Jenkins N.A.,
RA Copeland N.G., Gridley T.;
RA "Cloning, analysis, and chromosomal localization of Notch-1, a mouse
RT homolog of Drosophila Notch.";
RL Genomics 15:259-264(1993).
RN [2]
RP SEQUENCE OF 1551-2170 FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=93048835; PubMed=1425352;
RX Franco del Amo F., Smith D.E., Swiatek P.J., Gendron-Maguire M.,
RX Greenspan R.J., McMahon A.P., Gridley T.;
RT "Expression pattern of Notch, a mouse homolog of Drosophila Notch,
RT suggests an important role in early postimplantation mouse
RT development.";
RL Development 115:737-744(1992).
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED ALMOST UNIFORMLY IN EARLY EMBRYOS.
CC -1- SIMILARITY: CONTAINS 36 EGF-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 5 LIN/NOTCH REPEATS.
CC -1- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC -1- SIMILARITY: HIGH, WITH OTHER NOTCH-TYPE PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

```

CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Z11886; CAA77941.1; -
CC HSSP: P00740; 1IXA.
CC MGD: MGI:97363; Notchl.
CC InterPro: IPR000152; -
CC InterPro: IPR000561; -
CC InterPro: IPR000800; -
CC InterPro: IPR001438; -
CC InterPro: IPR001881; -
CC InterPro: IPR002110; -
CC Pfam: PF00008; EGF_35.
CC Pfam: PF00023; ank; 6.
CC Pfam: PF00066; notch; 3.
CC PRINTS: PR00010; EGFBL00D.
CC PROSITE: PS50088; ANK_REPEAT; 2.
CC PROSITE: PS50297; ANK_REPEAT_REGION; 1.
CC PROSITE: PS00010; ASX_HYDROXYL; 22.
CC PROSITE: PS00022; EGF_1; 34.
CC PROSITE: PS01186; EGF_2; 27.
CC PROSITE: PS01187; EGF_CA; 21.
CC Differentiation: Neurogenesis; Repeat; ANK repeat; EGF-like domain;
KW Transmembrane; Signal; Glycoprotein.
FT SIGNAL 1 18 POTENTIAL.
FT CHAIN 19 2531 NEUROGENIC LOCUS NOTCH HOMOLOG PROTEIN 1.
FT DOMAIN 19 1725 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1726 1746 POTENTIAL.
FT DOMAIN 1747 2531 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 20 58 EGF-LIKE 1.
FT DOMAIN 59 99 EGF-LIKE 2.
FT DOMAIN 102 139 EGF-LIKE 3.
FT DOMAIN 140 176 EGF-LIKE 4.
FT DOMAIN 178 216 EGF-LIKE 5.
FT DOMAIN 218 255 EGF-LIKE 6.
FT DOMAIN 257 293 EGF-LIKE 7.
FT DOMAIN 295 333 EGF-LIKE 8.
FT DOMAIN 335 371 EGF-LIKE 9.
FT DOMAIN 372 410 EGF-LIKE 10.
FT DOMAIN 412 450 EGF-LIKE 11.
FT DOMAIN 452 488 EGF-LIKE 12.
FT DOMAIN 490 526 EGF-LIKE 13.
FT DOMAIN 528 564 EGF-LIKE 14.
FT DOMAIN 566 601 EGF-LIKE 15.
FT DOMAIN 603 639 EGF-LIKE 16.
FT DOMAIN 641 676 EGF-LIKE 17.
FT DOMAIN 678 714 EGF-LIKE 18.
FT DOMAIN 716 751 EGF-LIKE 19.
FT DOMAIN 753 789 EGF-LIKE 20.
FT DOMAIN 791 827 EGF-LIKE 21.
FT DOMAIN 829 867 EGF-LIKE 22.
FT DOMAIN 869 905 EGF-LIKE 23.
FT DOMAIN 907 943 EGF-LIKE 24.
FT DOMAIN 945 981 EGF-LIKE 25.
FT DOMAIN 983 1019 EGF-LIKE 26.
FT DOMAIN 1021 1057 EGF-LIKE 27.
FT DOMAIN 1059 1095 EGF-LIKE 28.
FT DOMAIN 1097 1143 EGF-LIKE 29.
FT DOMAIN 1145 1181 EGF-LIKE 30.
FT DOMAIN 1183 1219 EGF-LIKE 31.
FT DOMAIN 1221 1265 EGF-LIKE 32.
FT DOMAIN 1267 1305 EGF-LIKE 33.
FT DOMAIN 1307 1346 EGF-LIKE 34.
FT DOMAIN 1348 1384 EGF-LIKE 35.
FT DOMAIN 1387 1426 EGF-LIKE 36.
FT DOMAIN 1449 1462 CYS-RICH.
FT REPEAT 1445 1480 LIN/NOTCH 1.
FT REPEAT 1481 1522 LIN/NOTCH 2.
FT REPEAT 1523 1562 LIN/NOTCH 3.
FT REPEAT 1917 1947 ANK 1.
FT REPEAT 1949 1979 ANK 2.
FT REPEAT 1983 2012 ANK 3.
FT REPEAT 2016 2045 ANK 4.
FT REPEAT 2049 2078 ANK 5.

```

```
FT DISULFID 24 BY SIMILARITY.
FT DISULFID 31 BY SIMILARITY.
FT DISULFID 63 BY SIMILARITY.
FT DISULFID 68 BY SIMILARITY.
FT DISULFID 86 BY SIMILARITY.
FT DISULFID 109 BY SIMILARITY.
FT DISULFID 111 BY SIMILARITY.
FT DISULFID 127 BY SIMILARITY.
FT DISULFID 129 BY SIMILARITY.
FT DISULFID 144 BY SIMILARITY.
FT DISULFID 149 BY SIMILARITY.
FT DISULFID 166 BY SIMILARITY.
FT DISULFID 182 BY SIMILARITY.
FT DISULFID 189 BY SIMILARITY.
FT DISULFID 206 BY SIMILARITY.
FT DISULFID 222 BY SIMILARITY.
FT DISULFID 227 BY SIMILARITY.
FT DISULFID 245 BY SIMILARITY.
FT DISULFID 261 BY SIMILARITY.
FT DISULFID 266 BY SIMILARITY.
FT DISULFID 283 BY SIMILARITY.
FT DISULFID 299 BY SIMILARITY.
FT DISULFID 306 BY SIMILARITY.
FT DISULFID 323 BY SIMILARITY.
FT DISULFID 339 BY SIMILARITY.
FT DISULFID 344 BY SIMILARITY.
FT DISULFID 361 BY SIMILARITY.
FT DISULFID 376 BY SIMILARITY.
FT DISULFID 381 BY SIMILARITY.
FT DISULFID 400 BY SIMILARITY.
FT DISULFID 416 BY SIMILARITY.
FT DISULFID 423 BY SIMILARITY.
FT DISULFID 440 BY SIMILARITY.
FT DISULFID 456 BY SIMILARITY.
FT DISULFID 461 BY SIMILARITY.
FT DISULFID 478 BY SIMILARITY.
FT DISULFID 494 BY SIMILARITY.
FT DISULFID 499 BY SIMILARITY.
FT DISULFID 516 BY SIMILARITY.
FT DISULFID 532 BY SIMILARITY.
FT DISULFID 537 BY SIMILARITY.
FT DISULFID 554 BY SIMILARITY.
FT DISULFID 570 BY SIMILARITY.
FT DISULFID 575 BY SIMILARITY.
FT DISULFID 591 BY SIMILARITY.
FT DISULFID 607 BY SIMILARITY.
FT DISULFID 612 BY SIMILARITY.
FT DISULFID 629 BY SIMILARITY.
FT DISULFID 645 BY SIMILARITY.
FT DISULFID 666 BY SIMILARITY.
FT DISULFID 682 BY SIMILARITY.
FT DISULFID 687 BY SIMILARITY.
FT DISULFID 704 BY SIMILARITY.
FT DISULFID 720 BY SIMILARITY.
FT DISULFID 725 BY SIMILARITY.
FT DISULFID 741 BY SIMILARITY.
FT DISULFID 757 BY SIMILARITY.
FT DISULFID 762 BY SIMILARITY.
FT DISULFID 779 BY SIMILARITY.
FT DISULFID 795 BY SIMILARITY.
FT DISULFID 800 BY SIMILARITY.
FT DISULFID 817 BY SIMILARITY.
FT DISULFID 833 BY SIMILARITY.
FT DISULFID 838 BY SIMILARITY.
FT DISULFID 857 BY SIMILARITY.
FT DISULFID 873 BY SIMILARITY.
FT DISULFID 878 BY SIMILARITY.
FT DISULFID 895 BY SIMILARITY.
FT DISULFID 911 BY SIMILARITY.
FT DISULFID 922 BY SIMILARITY.
FT DISULFID 933 BY SIMILARITY.
FT DISULFID 942 BY SIMILARITY.
FT DISULFID 957 BY SIMILARITY.
FT DISULFID 992 BY SIMILARITY.
FT DISULFID 1007 BY SIMILARITY.
FT DISULFID 1018 BY SIMILARITY.
FT DISULFID 1025 BY SIMILARITY.
FT DISULFID 1030 BY SIMILARITY.
FT DISULFID 1045 BY SIMILARITY.
FT DISULFID 1063 BY SIMILARITY.
FT DISULFID 1074 BY SIMILARITY.
FT DISULFID 1083 BY SIMILARITY.
FT DISULFID 1085 BY SIMILARITY.
FT DISULFID 1101 BY SIMILARITY.
FT DISULFID 1116 BY SIMILARITY.
FT DISULFID 1131 BY SIMILARITY.
FT DISULFID 1133 BY SIMILARITY.
FT DISULFID 1142 BY SIMILARITY.
FT DISULFID 1149 BY SIMILARITY.
FT DISULFID 1160 BY SIMILARITY.
Query Match 1.9%; Score 9; DB 1; Length 2531;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 82 PCHNGGTCE 90
Db 686 PCHNGGTCE 694
|||||||
RESULT 16
NTCL_RAT STANDARD; PRT; 2531 AA.
AC Q07008;
DT 01-NOV-1995 (Rel. 32, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE NEUROGENIC LOCUS NOTCH HOMOLOG PROTEIN 1 PRECURSOR.
GN NOTCH1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Schwann cell;
RX MEDLINE=92111383; PubMed=1764995;
RA Weinmaster G., Roberts V.J., Lemke G.;
RT "A homolog of Drosophila Notch expressed during mammalian
development.";
RL Development 113:199-205(1991).
RN [2]
RP REVISIONS TO 1652-1653.
RA Weinmaster G.;
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: REQUIRED FOR THE CORRECT DIFFERENTIATION OF A NUMBER
OF TISSUES.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- DEVELOPMENTAL STAGE: IN THE EMBRYO, HIGHEST LEVELS OCCUR BETWEEN
DAYS 12 AND 14 AND DECREASE RAPIDLY TO MUCH LOWER LEVELS IN THE
ADULT.
CC -!- SIMILARITY: HIGH, WITH OTHER NOTCH-TYPE PROTEINS.
CC -!- SIMILARITY: CONTAINS 36 EGF-LIKE DOMAINS.
CC -!- SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
CC -!- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X57405; CAA40667.1; -
DR HSSP; P00740; IIXA.
DR InterPro; IPR000152; -
DR InterPro; IPR000561; -
DR InterPro; IPR000800; -
DR InterPro; IPR001438; -
DR InterPro; IPR001881; -
DR InterPro; IPR002049; -
```

[illegible]

FT DISULFID 1154 1169 BY SIMILARITY.
 FT DISULFID 1171 1180 BY SIMILARITY.
 FT DISULFID 1187 1198 BY SIMILARITY.
 FT DISULFID 1192 1207 BY SIMILARITY.

Query Match 1.98; Score 9; DB 1; Length 2531;
 Best Local Similarity 100.0%; Pred. No. 1.1;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 82 PCHNGTGTC 90

DB 686 PCHNGTGTC 694

RESULT 17

DL_DROME STANDARD; PRT: 833 AA.
 AC P10041: Q9VDY2; Q99108;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE NEUROGENIC LOCUS DELTA PROTEIN PRECURSOR.
 GN DL OR CG3619.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A., TISSUE SPECIFICITY, AND DEVELOPMENTAL STAGE.
 RC TISSUE=Embryo;
 RA Vaessin H., Bremer K.A., Knust E., Campos-Ortega J.A.;
 RT "The neurogenic gene Delta of Drosophila melanogaster is expressed in
 neurogenic territories and encodes a putative transmembrane protein
 with EGF-like repeats";
 RL EMBO J. 6:3431-3440(1987).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=OREGON-R; TISSUE=Embryo;
 RX MEDLINE=89196890; PubMed=3149249;
 RA Kopczyński C.C., Alton A.K., Fechtel K., Kooh P.J., Muskavitch M.A.T.;
 RT "Delta, a Drosophila neurogenic gene, is transcriptionally complex and
 encodes a protein related to blood coagulation factors and epidermal
 growth factor of vertebrates";
 RL Genes Dev. 2:1723-1735(1988).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Ananides P.G., Scher S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champagne M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.D.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bertram B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Dou L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-F., Wassarman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster";
 RL Science 287:2185-2195(2000).
 RN [4]
 RP SEQUENCE OF 422-621 FROM N.A.
 RC TISSUE=Embryo;
 RX MEDLINE=87218537; PubMed=3107986;
 RA Knust E., Dietrich U., Tepass U., Bremer K.A., Weigel D., Vaessin H.,
 RA Campos-Ortega J.A.;
 RT "EGF homologous sequences encoded in the genome of Drosophila
 melanogaster, and their relation to neurogenic genes";
 RL EMBO J. 6:761-766(1987).
 RN [5]
 RP PATTERN OF TRANSCRIPTION, AND CHARACTERIZATION.
 RX MEDLINE=91209246; PubMed=2128477;
 RA Haenlin M., Kramatschek B., Campos-Ortega J.A.;
 RT "The pattern of transcription of the neurogenic gene Delta of
 Drosophila melanogaster";
 RL Development 110:905-914(1990).
 CC -!- FUNCTION: ESSENTIAL FOR PROPER DIFFERENTIATION OF ECTODERM. DL
 IS REQUIRED FOR THE CORRECT SEPARATION OF NEURAL AND EPIDERMAL
 CELL LINEAGES.
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -!- TISSUE SPECIFICITY: DETECTED IN ALL AREAS WITH NEUROGENIC
 ABILITIES, FOR EXAMPLE THE NEUROGENIC ECTODERM AND THE PRIMORDIA
 OF THE SENSE ORGANS. LATER EXPRESSION IS RESTRICTED TO THOSE CELLS
 THAT HAVE ADOPTED A NEURAL FATE.
 CC -!- DEVELOPMENTAL STAGE: EXPRESSED BOTH MATERNALLY AND ZYGOTICALLY.
 CC EXPRESSION IS HIGHEST EARLY IN EMBRYONIC DEVELOPMENT (STAGE 5) AND
 REDUCES TO A LOW LEVEL DURING LARVAL STAGES.
 CC -!- MISCELLANEOUS: SEPARATION OF NEUROBLASTS FROM THE ECTODERM INTO
 THE INNER PART OF EMBRYO IS ONE OF THE FIRST STEPS OF CNS
 DEVELOPMENT IN INSECTS, THIS PROCESS IS UNDER CONTROL OF THE
 NEUROGENIC GENES.
 CC -!- MISCELLANEOUS: NOTCH AND SERRATE MAY INTERACT AT THE PROTEIN
 LEVEL, IT IS CONCEIVABLE THAT THE SERRATE AND DELTA PROTEINS MAY
 COMPETE FOR BINDING WITH THE NOTCH PROTEIN.
 CC -!- SIMILARITY: CONTAINS 9 EGF-LIKE DOMAINS.
 CC -!- SIMILARITY: BELONGS TO THE DELTA/SERRATE/JAGGED FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X06289; CAA29617.1; -
 CC EMBL; Y00222; CAA68369.1; -
 CC EMBL; AE003725; AAF55657.1; -
 CC EMBL; X05140; CAA28786.1; -
 CC PIR; S00670; S00670.
 CC PIR; A26637; A26637.
 CC HSP; P00740; I1XA.
 CC FlyBase; FBgn0000463; DL.
 CC InterPro; IPR000152; -
 CC InterPro; IPR000561; -
 CC InterPro; IPR001774; -
 CC InterPro; IPR001881; -
 CC Pfam; PF01414; DSL; 1.
 CC Pfam; PF00008; EGF; 9.

DR PROSITE; PS00010; ASX-HYDROXYL; 3.
DR PROSITE; PS00022; EGF_1; 9.
DR PROSITE; PS01186; EGF_2; 9.
DR PROSITE; PS01187; EGF_CA; 2.
KW Differentiation; Neurogenesis; Repeat; Transmembrane;
EGF-like domain; Glycoprotein; Signal.
FT SIGNAL 1 18 POTENTIAL.
FT CHAIN 19 833 NEUROGENIC LOCUS DELTA PROTEIN.
FT DOMAIN 19 653 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 654 677 POTENTIAL.
FT DOMAIN 678 833 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 227 258 EGF-LIKE 1.
FT DOMAIN 256 289 EGF-LIKE 2.
FT DOMAIN 291 329 EGF-LIKE 3.
FT DOMAIN 331 372 EGF-LIKE 4.
FT DOMAIN 374 416 EGF-LIKE 5.
FT DOMAIN 418 451 EGF-LIKE 6.
FT DOMAIN 453 489 EGF-LIKE 7, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 491 527 EGF-LIKE 8.
FT DOMAIN 529 565 EGF-LIKE 9, CALCIUM-BINDING (POTENTIAL).
FT DISULFID 231 240 BY SIMILARITY.
FT DISULFID 235 246 BY SIMILARITY.
FT DISULFID 248 257 BY SIMILARITY.
FT DISULFID 260 271 BY SIMILARITY.
FT DISULFID 266 277 BY SIMILARITY.
FT DISULFID 279 288 BY SIMILARITY.
FT DISULFID 295 307 BY SIMILARITY.
FT DISULFID 301 317 BY SIMILARITY.
FT DISULFID 319 328 BY SIMILARITY.
FT DISULFID 335 348 BY SIMILARITY.
FT DISULFID 342 360 BY SIMILARITY.
FT DISULFID 362 371 BY SIMILARITY.
FT DISULFID 378 388 BY SIMILARITY.
FT DISULFID 383 404 BY SIMILARITY.
FT DISULFID 406 415 BY SIMILARITY.
FT DISULFID 422 433 BY SIMILARITY.
FT DISULFID 427 439 BY SIMILARITY.
FT DISULFID 441 450 BY SIMILARITY.
FT DISULFID 457 468 BY SIMILARITY.
FT DISULFID 462 477 BY SIMILARITY.
FT DISULFID 479 488 BY SIMILARITY.
FT DISULFID 495 506 BY SIMILARITY.
FT DISULFID 500 515 BY SIMILARITY.
FT DISULFID 517 526 BY SIMILARITY.
FT DISULFID 533 544 BY SIMILARITY.
FT DISULFID 538 553 BY SIMILARITY.
FT DISULFID 555 564 BY SIMILARITY.
FT CARBOHYD 98 98 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 137 137 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 167 167 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 649 649 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 5 5 K -> N (IN REF. 1).
FT CONFLICT 67 67 V -> L (IN REF. 1).
FT CONFLICT 363 363 A -> R (IN REF. 1).
FT CONFLICT 437 438 GK -> ET (IN REF. 3).
FT CONFLICT 443 443 A -> S (IN REF. 1 AND 3).
FT CONFLICT 459 459 G -> A (IN REF. 3).
FT CONFLICT 490 490 S -> T (IN REF. 3).
FT CONFLICT 591 591 T -> A (IN REF. 1 AND 3).
FT CONFLICT 631 631 D -> N (IN REF. 1).
FT CONFLICT 652 652 G -> A (IN REF. 1).
FT CONFLICT 662 662 L -> M (IN REF. 1).
FT CONFLICT 759 759 S -> T (IN REF. 1).
FT CONFLICT 801 804 ACSS -> DLFI (IN REF. 1).
FT CONFLICT 812 812 V -> A (IN REF. 2).
FT CONFLICT 815 817 SGA -> TD (IN REF. 1).
FT CONFLICT 833 833 M -> YVTPKIRKGSWEIRNPHGGADSTYTKRLGLWQNV
RETPKMLLIEAV (IN REF. 1).
SQ SEQUENCE 833 AA; 88840 MW; EDBA10A70003D9A1 CRC64;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 82 PCHNGGTC 89
DB 537 PCHNGGTC 544
RESULT 18
PCGB_MOUSE
ID PCGB_MOUSE STANDARD; PRT; 883 AA.
AC Q61361;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE BREVICAN CORE PROTEIN PRECURSOR.
GN BCAN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BALB/C; TISSUE=Brain;
RX MEDLINE=97432816; PubMed=9286696;
RA Rauch U., Meyer H., Brakebusch C., Seidenbecher C., Gundelfinger E.D.,
Beier D.R., Fassler R.;
RT "Sequence and chromosomal localization of the mouse brevicane gene";
RL Genomics 44:15-21(1997).
CC -!- FUNCTION: MAY PLAY A ROLE IN THE TERMINALLY DIFFERENTIATING AND
THE ADULT NERVOUS SYSTEM DURING POSTNATAL DEVELOPMENT. COULD
STABILIZE INTERACTIONS BETWEEN HA AND BRAIN PROTEOGLYCAN.
CC -!- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR MATRIX (BY
SIMILARITY).
CC -!- TISSUE SPECIFICITY: BRAIN (BY SIMILARITY).
CC -!- PTM: CONTAINS MOSTLY CHONDROITIN SULFATE (BY SIMILARITY).
CC -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.
CC -!- SIMILARITY: CONTAINS 2 LINK DOMAINS.
CC -!- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 SUSHI (SCR) REPEAT.
CC -!- SIMILARITY: BELONGS TO THE AGGREGAN/VERSICAN PROTEOGLYCAN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC EMBL; X87096; CAA60575.1; .
CC HSP; P20693; 1HLJ.
CC MGD; MGI:1096385; Bcan.
CC InterPro; IPR000436; .
CC InterPro; IPR000495; .
CC InterPro; IPR000538; .
CC InterPro; IPR000561; .
CC InterPro; IPR001304; .
CC InterPro; IPR003006; .
CC Pfam; PF00008; EGF; 1.
CC Pfam; PF00193; Xlink; 2.
CC Pfam; PF00047; ig; 1.
CC Pfam; PF00059; lectin_c; 1.
CC Pfam; PF00084; sush; 1.
CC PROSITE; PS00022; EGF_1; 1.
CC PROSITE; PS01186; EGF_2; 1.
CC PROSITE; PS00290; IG_MHC; 1.
CC PROSITE; PS01241; LINK; 2.
CC PROSITE; PS00615; C-TYPE LECTIN_1; 1.
CC PROSITE; PS00041; C-TYPE LECTIN_2; 1.
KW Glycoprotein; Hyaluronic acid; Proteoglycan; Lectin; Signal; Sush; EGF-like domain; Repeat; Immunoglobulin domain.
FT SIGNAL 1 22 POTENTIAL.

Query Match 1.78; Score 8; DB 1; Length 833;
Best Local Similarity 100.0%; Pred. No. 4.5;

```

FT CHAIN 23 883 BREVICAN CORE PROTEIN.
FT DOMAIN 32 157 IG-LIKE V-TYPE DOMAIN.
FT DOMAIN 173 250 LINK 1.
FT DOMAIN 271 352 LINK 2.
FT DOMAIN 622 658 EGF-LIKE.
FT DOMAIN 658 786 C-TYPE LECTIN.
FT DOMAIN 787 851 SUSHI.
FT DISULFID 56 136 BY SIMILARITY.
FT DISULFID 178 249 BY SIMILARITY.
FT DISULFID 202 223 BY SIMILARITY.
FT DISULFID 276 351 BY SIMILARITY.
FT DISULFID 300 321 BY SIMILARITY.
FT DISULFID 626 637 BY SIMILARITY.
FT DISULFID 631 646 BY SIMILARITY.
FT DISULFID 648 657 BY SIMILARITY.
FT DISULFID 664 675 BY SIMILARITY.
FT DISULFID 692 784 BY SIMILARITY.
FT DISULFID 760 776 BY SIMILARITY.
FT DISULFID 791 834 BY SIMILARITY.
FT DISULFID 820 847 BY SIMILARITY.
FT CARBOHYD 129 129 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 336 336 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 883 AA; 96013 MW; CC2C33C97B453E45 CRC64;

Query Match 1.7%; Score 8; DB 1; Length 883;
Best Local Similarity 100.0%; Pred. No. 4.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 82 PCHNGGTC 89
Db 630 PCHNGGTC 637
|||||||

RESULT 19
PGCB_RAT STANDARD; PRT; 883 AA.
AC P5068; Q63040; Q62860; Q63513;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE BREVICAN CORE PROTEIN PRECURSOR (BRAIN ENRICHED HYALURONAN BINDING
DE PROTEIN) (BEHAB PROTEIN).
GN BCAN.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
[1]
SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Brain;
RX MEDLINE=96070828; PubMed=7592978;
RA Seidenbecher C.I., Richter K., Rauch U., Faessler R., Garner C.C.,
RA Gundelfinger E.D.;
RT "Brevican, a chondroitin sulfate proteoglycan of rat brain, occurs as
RT secreted and cell surface glycosylphosphatidylinositol-anchored
RT isoforms.";
RL J. Biol. Chem. 270:27206-27212(1995).
[2]
SEQUENCE FROM N.A., AND SEQUENCE OF 396-407.
RN
RP
RC TISSUE=Brain;
RX MEDLINE=96074575; PubMed=7488217;
RA Yamada H., Watanabe K., Shimonaka M., Yamasaki M., Yamaguchi Y.;
RT "cDNA cloning and the identification of an aggrecanase-like cleavage
RL site in rat brevican.";
RN Biochem. Biophys. Res. Commun. 216:957-963(1995).
[3]
SEQUENCE OF 1-423 FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Brain;
RX MEDLINE=94216386; PubMed=7512973;
RA Jaworski D.M., Kelly G.M., Hockfield S.;
RT "BEHAB, a new member of the proteoglycan tandem repeat family of
RT hyaluronan-binding proteins that is restricted to the brain.";

```

```

RL J. Cell Biol. 125:495-509(1994).
CC !- FUNCTION: MAY PLAY A ROLE IN THE TERMINALLY DIFFERENTIATING AND
CC THE ADULT NERVOUS SYSTEM DURING POSTNATAL DEVELOPMENT. COULD
CC STABILIZE INTERACTIONS BETWEEN HA AND BRAIN PROTEOGLYCAN. THE
CC GPI-ANCHORED ISOFORM MAY FUNCTION AS A CHONDROITIN SULFATE-
CC BEARING CELL SURFACE RECEPTOR.
CC !- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR MATRIX AND ONE FORM
CC ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR.
CC !- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A SECRETED FORM (SHOWN HERE) AND
CC A GPI-ANCHORED FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC !- TISSUE SPECIFICITY: BRAIN.
CC !- DEVELOPMENTAL STAGE: SOLUBLE FORM INCREASES FROM DAY P4 TO P64.
CC GPI-ANCHORED ISOFORM INCREASES AFTER DAY P8.
CC !- PTM: CONTAINS MOSTLY CHONDROITIN SULFATE.
CC !- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.
CC !- SIMILARITY: CONTAINS 2 LINK DOMAINS.
CC !- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC !- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC !- SIMILARITY: CONTAINS 1 SUSHI (SCR) REPEAT.
CC !- SIMILARITY: BELONGS TO THE AGGRECAN/VERSICAN PROTEOGLYCAN FAMILY.
CC !- CAUTION: REF.3 SEQUENCE DIFFERS FROM THAT SHOWN FROM POSITION 364
CC ONWARD AND IS SMALLER (371 AA) DUE TO A FRAMESHIFT.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X79881; CAA56255.1; -
CC EMBL; X86406; CAA60160.1; -
CC EMBL; U37142; AAA87847.1; -
CC EMBL; Z28366; CAA82215.1; ALT_FRAME.
CC HSSP; P20693; IHLJ.
CC InterPro; IPR000436; -
CC InterPro; IPR000495; -
CC InterPro; IPR000538; -
CC InterPro; IPR000561; -
CC InterPro; IPR001304; -
CC InterPro; IPR003006; -
CC Pfam; PF00008; EGF_1; 1.
CC Pfam; PF00193; Xlink; 2.
CC Pfam; PF00047; ig_1; 1.
CC Pfam; PF00059; lectin_C; 1.
CC ProSITE; PS00022; EGF_1; 1.
CC ProSITE; PS01186; EGF_2; 1.
CC ProSITE; PS00290; IG_MHC; 1.
CC ProSITE; PS01241; LINK; 2.
CC ProSITE; PS00615; C-TYPE_LECTIN_1; 1.
CC ProSITE; PS00615; C-TYPE_LECTIN_2; 1.
KW Glycoprotein; Hyaluronic acid; Proteoglycan; Lectin; Signal; Sushi;
KW EGF-like domain; Repeat; Immunoglobulin domain; Alternative splicing;
KW GPI-anchor.
FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 883 BREVICAN CORE PROTEIN.
FT DOMAIN 32 157 IG-LIKE V-TYPE DOMAIN.
FT DOMAIN 173 250 LINK 1.
FT DOMAIN 271 352 LINK 2.
FT DOMAIN 622 658 EGF-LIKE.
FT DOMAIN 658 786 C-TYPE LECTIN.
FT DOMAIN 787 851 SUSHI.
FT DISULFID 56 136 BY SIMILARITY.
FT DISULFID 178 249 BY SIMILARITY.
FT DISULFID 202 223 BY SIMILARITY.
FT DISULFID 276 351 BY SIMILARITY.
FT DISULFID 300 321 BY SIMILARITY.
FT DISULFID 626 637 BY SIMILARITY.
FT DISULFID 631 646 BY SIMILARITY.
FT DISULFID 648 657 BY SIMILARITY.
FT DISULFID 791 834 BY SIMILARITY.

```

or send an email to license@isb-sib.ch.

CC	EMBL; X75887; CAA53481.1; -			
CC	HSSP; P20693; 1HLJ			
DR	InterPro; IPR000436; -			
DR	InterPro; IPR000495; -			
DR	InterPro; IPR000538; -			
DR	InterPro; IPR000561; -			
DR	InterPro; IPR001304; -			
DR	InterPro; IPR003006; -			
DR	Pfam; PF00008; EGF; 1.			
DR	Pfam; PF00193; Xlink; 2.			
DR	Pfam; PF00047; Ig; 1.			
DR	Pfam; PF00059; lectin_c; 1.			
DR	Pfam; PF00084; sushi; 1.			
DR	PROSITE; PS00022; EGF_1; 1.			
DR	PROSITE; PS01186; EGF_2; 1.			
DR	PROSITE; PS00290; IG_MHC; 1.			
DR	PROSITE; PS01241; LINK; 2.			
DR	PROSITE; PS00615; C-TYPE_LECTIN_1; 1.			
DR	PROSITE; PS00041; C-TYPE_LECTIN_2; 1.			
KW	Glycoprotein; Hyaluronic acid; Proteoglycan; Lectin; Signal; Sushi;			
KW	EGF-like domain; Repeat; Immunoglobulin domain.			
FT	SIGNAL	1	22	POTENTIAL.
FT	CHAIN	23	912	BREVICAN CORE PROTEIN.
FT	DOMAIN	32	158	IG-LIKE V-TYPE DOMAIN.
FT	DOMAIN	174	251	LINK 1.
FT	DOMAIN	272	353	LINK 2.
FT	DOMAIN	647	683	EGF-LIKE.
FT	DOMAIN	683	811	C-TYPE LECTIN.
FT	DOMAIN	812	876	SUSHI.
FT	DISULFID	57	137	BY SIMILARITY.
FT	DISULFID	179	250	BY SIMILARITY.
FT	DISULFID	203	224	BY SIMILARITY.
FT	DISULFID	277	352	BY SIMILARITY.
FT	DISULFID	301	322	BY SIMILARITY.
FT	DISULFID	651	662	BY SIMILARITY.
FT	DISULFID	656	671	BY SIMILARITY.
FT	DISULFID	673	682	BY SIMILARITY.
FT	DISULFID	689	700	BY SIMILARITY.
FT	DISULFID	717	809	BY SIMILARITY.
FT	DISULFID	785	801	BY SIMILARITY.
FT	DISULFID	816	859	BY SIMILARITY.
FT	DISULFID	845	872	BY SIMILARITY.
FT	CARBOHYD	130	130	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	337	337	N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ	SEQUENCE	912 AA;	99554 MW;	677B3EB1C688C4D7 CRC64;
Query Match 1.7%; Score 8; DB 1; Length 912;				
Best Local Similarity 100.0%; Pred. No. 4.9;				
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
QY	82	PCHNGGTC	89	
Db	655	PCHNGGTC	662	
RESULT 21				
NOTC_XENLA	STANDARD;	PRT;	2524 AA.	
ID	NOTC_XENLA	STANDARD;	PRT;	2524 AA.
AC	P21783;			
DT	01-MAY-1991 (Rel. 18, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	01-OCT-2000 (Rel. 40, Last annotation update)			
DE	NEUROGENIC LOCUS NOTCH PROTEIN HOMOLOG PRECURSOR (NOTCH PROTEIN).			
GN	XOTCH.			
OS	Xenopus laevis (African clawed frog).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;			
OC	Xenopodinae; Xenopus.			
OX	NCBI_TaxID=8355;			
RN	[1]			

SEQUENCE FROM N.A.
 RX MEDLINE=90385285; PubMed=2402639;
 RA Coffman C., Harris W., Kintner C.;
 RT "Notch, the xenopus homolog of Drosophila notch."; Science 249:1438-1441(1990).
 RL [2]
 RN REVISIONS TO 1759-1782.
 RA Kintner C.;
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED ALMOST UNIFORMLY IN EARLY EMBRYOS.
 CC -1- SIMILARITY: HIGH, WITH OTHER NOTCH-TYPE PROTEINS.
 CC -1- SIMILARITY: CONTAINS 36 EGF-LIKE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
 CC -1- SIMILARITY: CONTAINS 6 ANK REPEATS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M33874; AAB02039.1; -
 DR PIR: A35844; A35844.
 DR RSP: P00740; IIXA.
 DR InterPro: IPR000152; -
 DR InterPro: IPR000561; -
 DR InterPro: IPR000800; -
 DR InterPro: IPR001438; -
 DR InterPro: IPR001881; -
 DR InterPro: IPR002110; -
 DR Pfam: PF00008; EGF_36.
 DR Pfam: PF00023; ank; 6.
 DR Pfam: PF00066; notch; 3.
 DR PRINTS: PR00010; EGFBLD.
 DR PROSITE: PS50088; ANK_REPEAT; 4.
 DR PROSITE: PS0297; ANK_REPEAT_REGION; 1.
 DR PROSITE: PS00010; ASX_HYDROXYL; 23.
 DR PROSITE: PS00022; EGF_1; 34.
 DR PROSITE: PS01186; EGF_2; 29.
 DR PROSITE: PS01187; EGF_CA; 21.
 DR PROSITE: PS01187; Neurogenesis; Repeat; ANK repeat; EGF-like domain;
 KW Transmembrane; Signal; Glycoprotein.
 FT SIGNAL 1 19
 FT CHAIN 20 2524
 FT DOMAIN 20 1728
 FT TRANSMEM 1729 1750
 FT DOMAIN 1751 2524
 FT DOMAIN 20 57
 FT DOMAIN 58 99
 FT DOMAIN 102 140
 FT DOMAIN 141 177
 FT DOMAIN 179 215
 FT DOMAIN 217 254
 FT DOMAIN 256 292
 FT DOMAIN 294 332
 FT DOMAIN 334 370
 FT DOMAIN 371 409
 FT DOMAIN 411 449
 FT DOMAIN 451 487
 FT DOMAIN 489 525
 FT DOMAIN 527 563
 FT DOMAIN 565 600
 FT DOMAIN 602 638
 FT DOMAIN 640 675
 FT DOMAIN 677 713
 FT DOMAIN 715 750
 FT DOMAIN 752 788
 FT DOMAIN 790 826
 FT DOMAIN 828 866
 FT DOMAIN 868 904

FT DOMAIN 906
 FT DOMAIN 944
 FT DOMAIN 982
 FT DOMAIN 1018
 FT DOMAIN 1020
 FT DOMAIN 1056
 FT DOMAIN 1094
 FT DOMAIN 1096
 FT DOMAIN 1140
 FT DOMAIN 1142
 FT DOMAIN 1180
 FT DOMAIN 1182
 FT DOMAIN 1218
 FT DOMAIN 1220
 FT DOMAIN 1264
 FT DOMAIN 1304
 FT DOMAIN 1346
 FT DOMAIN 1366
 FT DOMAIN 1377
 FT DOMAIN 1383
 FT DOMAIN 1424
 FT DOMAIN 1436
 FT REPEAT 1441
 FT REPEAT 1479
 FT REPEAT 1520
 FT REPEAT 1521
 FT REPEAT 1560
 FT REPEAT 1562
 FT REPEAT 1566
 FT REPEAT 1570
 FT REPEAT 1574
 FT REPEAT 1578
 FT REPEAT 1582
 FT REPEAT 1586
 FT REPEAT 1590
 FT REPEAT 1594
 FT REPEAT 1598
 FT REPEAT 1602
 FT REPEAT 1606
 FT REPEAT 1610
 FT REPEAT 1614
 FT REPEAT 1618
 FT REPEAT 1622
 FT REPEAT 1626
 FT REPEAT 1630
 FT REPEAT 1634
 FT REPEAT 1638
 FT REPEAT 1642
 FT REPEAT 1646
 FT REPEAT 1650
 FT REPEAT 1654
 FT REPEAT 1658
 FT REPEAT 1662
 FT REPEAT 1666
 FT REPEAT 1670
 FT REPEAT 1674
 FT REPEAT 1678
 FT REPEAT 1682
 FT REPEAT 1686
 FT REPEAT 1690
 FT REPEAT 1694
 FT REPEAT 1698
 FT REPEAT 1702
 FT REPEAT 1706
 FT REPEAT 1710
 FT REPEAT 1714
 FT REPEAT 1718
 FT REPEAT 1722
 FT REPEAT 1726
 FT REPEAT 1730
 FT REPEAT 1734
 FT REPEAT 1738
 FT REPEAT 1742
 FT REPEAT 1746
 FT REPEAT 1750
 FT REPEAT 1754
 FT REPEAT 1758
 FT REPEAT 1762
 FT REPEAT 1766
 FT REPEAT 1770
 FT REPEAT 1774
 FT REPEAT 1778
 FT REPEAT 1782
 FT REPEAT 1786
 FT REPEAT 1790
 FT REPEAT 1794
 FT REPEAT 1798
 FT REPEAT 1802
 FT REPEAT 1806
 FT REPEAT 1810
 FT REPEAT 1814
 FT REPEAT 1818
 FT REPEAT 1822
 FT REPEAT 1826
 FT REPEAT 1830
 FT REPEAT 1834
 FT REPEAT 1838
 FT REPEAT 1842
 FT REPEAT 1846
 FT REPEAT 1850
 FT REPEAT 1854
 FT REPEAT 1858
 FT REPEAT 1862
 FT REPEAT 1866
 FT REPEAT 1870
 FT REPEAT 1874
 FT REPEAT 1878
 FT REPEAT 1882
 FT REPEAT 1886
 FT REPEAT 1890
 FT REPEAT 1894
 FT REPEAT 1898
 FT REPEAT 1902
 FT REPEAT 1906
 FT REPEAT 1910
 FT REPEAT 1914
 FT REPEAT 1918
 FT REPEAT 1922
 FT REPEAT 1926
 FT REPEAT 1930
 FT REPEAT 1934
 FT REPEAT 1938
 FT REPEAT 1942
 FT REPEAT 1946
 FT REPEAT 1950
 FT REPEAT 1954
 FT REPEAT 1958
 FT REPEAT 1962
 FT REPEAT 1966
 FT REPEAT 1970
 FT REPEAT 1974
 FT REPEAT 1978
 FT REPEAT 1982
 FT REPEAT 1986
 FT REPEAT 1990
 FT REPEAT 1994
 FT REPEAT 1998
 FT REPEAT 2002
 FT REPEAT 2006
 FT REPEAT 2010
 FT REPEAT 2014
 FT REPEAT 2018
 FT REPEAT 2022
 FT REPEAT 2026
 FT REPEAT 2030
 FT REPEAT 2034
 FT REPEAT 2038
 FT REPEAT 2042
 FT REPEAT 2046
 FT REPEAT 2050
 FT REPEAT 2054
 FT REPEAT 2058
 FT REPEAT 2062
 FT REPEAT 2066
 FT REPEAT 2070
 FT REPEAT 2074
 FT REPEAT 2078
 FT REPEAT 2082
 FT REPEAT 2086
 FT REPEAT 2090
 FT REPEAT 2094
 FT REPEAT 2098
 FT REPEAT 2102
 FT REPEAT 2106
 FT REPEAT 2110
 FT REPEAT 2114
 FT REPEAT 2118
 FT REPEAT 2122
 FT REPEAT 2126
 FT REPEAT 2130
 FT REPEAT 2134
 FT REPEAT 2138
 FT REPEAT 2142
 FT REPEAT 2146
 FT REPEAT 2150
 FT REPEAT 2154
 FT REPEAT 2158
 FT REPEAT 2162
 FT REPEAT 2166
 FT REPEAT 2170
 FT REPEAT 2174
 FT REPEAT 2178
 FT REPEAT 2182
 FT REPEAT 2186
 FT REPEAT 2190
 FT REPEAT 2194
 FT REPEAT 2198
 FT REPEAT 2202
 FT REPEAT 2206
 FT REPEAT 2210
 FT REPEAT 2214
 FT REPEAT 2218
 FT REPEAT 2222
 FT REPEAT 2226
 FT REPEAT 2230
 FT REPEAT 2234
 FT REPEAT 2238
 FT REPEAT 2242
 FT REPEAT 2246
 FT REPEAT 2250
 FT REPEAT 2254
 FT REPEAT 2258
 FT REPEAT 2262
 FT REPEAT 2266
 FT REPEAT 2270
 FT REPEAT 2274
 FT REPEAT 2278
 FT REPEAT 2282
 FT REPEAT 2286
 FT REPEAT 2290
 FT REPEAT 2294
 FT REPEAT 2298
 FT REPEAT 2302
 FT REPEAT 2306
 FT REPEAT 2310
 FT REPEAT 2314
 FT REPEAT 2318
 FT REPEAT 2322
 FT REPEAT 2326
 FT REPEAT 2330
 FT REPEAT 2334
 FT REPEAT 2338
 FT REPEAT 2342
 FT REPEAT 2346
 FT REPEAT 2350
 FT REPEAT 2354
 FT REPEAT 2358
 FT REPEAT 2362
 FT REPEAT 2366
 FT REPEAT 2370
 FT REPEAT 2374
 FT REPEAT 2378
 FT REPEAT 2382
 FT REPEAT 2386
 FT REPEAT 2390
 FT REPEAT 2394
 FT REPEAT 2398
 FT REPEAT 2402
 FT REPEAT 2406
 FT REPEAT 2410
 FT REPEAT 2414
 FT REPEAT 2418
 FT REPEAT 2422
 FT REPEAT 2426
 FT REPEAT 2430
 FT REPEAT 2434
 FT REPEAT 2438
 FT REPEAT 2442
 FT REPEAT 2446
 FT REPEAT 2450
 FT REPEAT 2454
 FT REPEAT 2458
 FT REPEAT 2462
 FT REPEAT 2466
 FT REPEAT 2470
 FT REPEAT 2474
 FT REPEAT 2478
 FT REPEAT 2482
 FT REPEAT 2486
 FT REPEAT 2490
 FT REPEAT 2494
 FT REPEAT 2498
 FT REPEAT 2502
 FT REPEAT 2506
 FT REPEAT 2510
 FT REPEAT 2514
 FT REPEAT 2518
 FT REPEAT 2522
 FT REPEAT 2526
 FT REPEAT 2530
 FT REPEAT 2534
 FT REPEAT 2538
 FT REPEAT 2542
 FT REPEAT 2546
 FT REPEAT 2550
 FT REPEAT 2554
 FT REPEAT 2558
 FT REPEAT 2562
 FT REPEAT 2566
 FT REPEAT 2570
 FT REPEAT 2574
 FT REPEAT 2578
 FT REPEAT 2582
 FT REPEAT 2586
 FT REPEAT 2590
 FT REPEAT 2594
 FT REPEAT 2598
 FT REPEAT 2602
 FT REPEAT 2606
 FT REPEAT 2610
 FT REPEAT 2614
 FT REPEAT 2618
 FT REPEAT 2622
 FT REPEAT 2626
 FT REPEAT 2630
 FT REPEAT 2634
 FT REPEAT 2638
 FT REPEAT 2642
 FT REPEAT 2646
 FT REPEAT 2650
 FT REPEAT 2654
 FT REPEAT 2658
 FT REPEAT 2662
 FT REPEAT 2666
 FT REPEAT 2670
 FT REPEAT 2674
 FT REPEAT 2678
 FT REPEAT 2682
 FT REPEAT 2686
 FT REPEAT 2690
 FT REPEAT 2694
 FT REPEAT 2698
 FT REPEAT 2702
 FT REPEAT 2706
 FT REPEAT 2710
 FT REPEAT 2714
 FT REPEAT 2718
 FT REPEAT 2722
 FT REPEAT 2726
 FT REPEAT 2730
 FT REPEAT 2734
 FT REPEAT 2738
 FT REPEAT 2742
 FT REPEAT 2746
 FT REPEAT 2750
 FT REPEAT 2754
 FT REPEAT 2758
 FT REPEAT 2762
 FT REPEAT 2766
 FT REPEAT 2770
 FT REPEAT 2774
 FT REPEAT 2778
 FT REPEAT 2782
 FT REPEAT 2786
 FT REPEAT 2790
 FT REPEAT 2794
 FT REPEAT 2798
 FT REPEAT 2802
 FT REPEAT 2806
 FT REPEAT 2810
 FT REPEAT 2814
 FT REPEAT 2818
 FT REPEAT 2822
 FT REPEAT 2826
 FT REPEAT 2830
 FT REPEAT 2834
 FT REPEAT 2838
 FT REPEAT 2842
 FT REPEAT 2846
 FT REPEAT 2850
 FT REPEAT 2854
 FT REPEAT 2858
 FT REPEAT 2862
 FT REPEAT 2866
 FT REPEAT 2870
 FT REPEAT 2874
 FT REPEAT 2878
 FT REPEAT 2882
 FT REPEAT 2886
 FT REPEAT 2890
 FT REPEAT 2894
 FT REPEAT 2898
 FT REPEAT 2902
 FT REPEAT 2906
 FT REPEAT 2910
 FT REPEAT 2914
 FT REPEAT 2918
 FT REPEAT 2922
 FT REPEAT 2926
 FT REPEAT 2930
 FT REPEAT 2934
 FT REPEAT 2938
 FT REPEAT 2942
 FT REPEAT 2946
 FT REPEAT 2950
 FT REPEAT 2954
 FT REPEAT 2958
 FT REPEAT 2962
 FT REPEAT 2966
 FT REPEAT 2970
 FT REPEAT 2974
 FT REPEAT 2978
 FT REPEAT 2982
 FT REPEAT 2986
 FT REPEAT 2990
 FT REPEAT 2994
 FT REPEAT 2998
 FT REPEAT 3002
 FT REPEAT 3006
 FT REPEAT 3010
 FT REPEAT 3014
 FT REPEAT 3018
 FT REPEAT 3022
 FT REPEAT 3026
 FT REPEAT 3030
 FT REPEAT 3034
 FT REPEAT 3038
 FT REPEAT 3042
 FT REPEAT 3046
 FT REPEAT 3050
 FT REPEAT 3054
 FT REPEAT 3058
 FT REPEAT 3062
 FT REPEAT 3066
 FT REPEAT 3070
 FT REPEAT 3074
 FT REPEAT 3078
 FT REPEAT 3082
 FT REPEAT 3086
 FT REPEAT 3090
 FT REPEAT 3094
 FT REPEAT 3098
 FT REPEAT 3102
 FT REPEAT 3106
 FT REPEAT 3110
 FT REPEAT 3114
 FT REPEAT 3118
 FT REPEAT 3122
 FT REPEAT 3126
 FT REPEAT 3130
 FT REPEAT 3134
 FT REPEAT 3138
 FT REPEAT 3142
 FT REPEAT 3146
 FT REPEAT 3150
 FT REPEAT 3154
 FT REPEAT 3158
 FT REPEAT 3162
 FT REPEAT 3166
 FT REPEAT 3170
 FT REPEAT 3174
 FT REPEAT 3178
 FT REPEAT 3182
 FT REPEAT 3186
 FT REPEAT 3190
 FT REPEAT 3194
 FT REPEAT 3198
 FT REPEAT 3202
 FT REPEAT 3206
 FT REPEAT 3210
 FT REPEAT 3214
 FT REPEAT 3218
 FT REPEAT 3222
 FT REPEAT 3226
 FT REPEAT 3230
 FT REPEAT 3234
 FT REPEAT 3238
 FT REPEAT 3242
 FT REPEAT 3246
 FT REPEAT 3250
 FT REPEAT 3254
 FT REPEAT 3258
 FT REPEAT 3262
 FT REPEAT 3266
 FT REPEAT 3270
 FT REPEAT 3274
 FT REPEAT 3278
 FT REPEAT 3282
 FT REPEAT 3286
 FT REPEAT 3290
 FT REPEAT 3294
 FT REPEAT 3298
 FT REPEAT 3302
 FT REPEAT 3306
 FT REPEAT 3310
 FT REPEAT 3314
 FT REPEAT 3318
 FT REPEAT 3322
 FT REPEAT 3326
 FT REPEAT 3330
 FT REPEAT 3334
 FT REPEAT 3338
 FT REPEAT 3342
 FT REPEAT 3346
 FT REPEAT 3350
 FT REPEAT 3354
 FT REPEAT 3358
 FT REPEAT 3362
 FT REPEAT 3366
 FT REPEAT 3370
 FT REPEAT 3374
 FT REPEAT 3378
 FT REPEAT 3382
 FT REPEAT 3386
 FT REPEAT 3390
 FT REPEAT 3394
 FT REPEAT 3398
 FT REPEAT 3402
 FT REPEAT 3406
 FT REPEAT 3410
 FT REPEAT 3414
 FT REPEAT 3418
 FT REPEAT 3422
 FT REPEAT 3426
 FT REPEAT 3430
 FT REPEAT 3434
 FT REPEAT 3438
 FT REPEAT 3442
 FT REPEAT 3446
 FT REPEAT 3450
 FT REPEAT 3454
 FT REPEAT 3458
 FT REPEAT 3462
 FT REPEAT 3466
 FT REPEAT 3470
 FT REPEAT 3474
 FT REPEAT 3478
 FT REPEAT 3482
 FT REPEAT 3486
 FT REPEAT 3490
 FT REPEAT 3494
 FT REPEAT 3498
 FT REPEAT 3502
 FT REPEAT 3506
 FT REPEAT 3510
 FT REPEAT 3514
 FT REPEAT 3518
 FT REPEAT 3522
 FT REPEAT 3526
 FT REPEAT 3530
 FT REPEAT 3534
 FT REPEAT 3538
 FT REPEAT 3542
 FT REPEAT 3546
 FT REPEAT 3550
 FT REPEAT 3554
 FT REPEAT 3558
 FT REPEAT 3562
 FT REPEAT 3566
 FT REPEAT 3570
 FT REPEAT 3574
 FT REPEAT 3578
 FT REPEAT 3582
 FT REPEAT 3586
 FT REPEAT 3590
 FT REPEAT 3594
 FT REPEAT 3598
 FT REPEAT 3602
 FT REPEAT 3606
 FT REPEAT 3610
 FT REPEAT 3614
 FT REPEAT 3618
 FT REPEAT 3622
 FT REPEAT 3626
 FT REPEAT 3630
 FT REPEAT 3634
 FT REPEAT 3638
 FT REPEAT 3642
 FT REPEAT 3646
 FT REPEAT 3650
 FT REPEAT 3654
 FT REPEAT 3658
 FT REPEAT 3662
 FT REPEAT 3666
 FT REPEAT 3670
 FT REPEAT 3674
 FT REPEAT 3678
 FT REPEAT 3682
 FT REPEAT 3686
 FT REPEAT 3690
 FT REPEAT 3694
 FT REPEAT 3698
 FT REPEAT 3702
 FT REPEAT 3706
 FT REPEAT 3710
 FT REPEAT 3714
 FT REPEAT 3718
 FT REPEAT 3722
 FT REPEAT 3726
 FT REPEAT 3730
 FT REPEAT 3734
 FT REPEAT 3738
 FT REPEAT 3742
 FT REPEAT 3746
 FT REPEAT 3750
 FT REPEAT 3754
 FT REPEAT 3758
 FT REPEAT 3762
 FT REPEAT 3766
 FT REPEAT 3770
 FT REPEAT 3774
 FT REPEAT 3778
 FT REPEAT 3782
 FT REPEAT 3786
 FT REPEAT 3790
 FT REPEAT 3794
 FT REPEAT 3798
 FT REPEAT 3802
 FT REPEAT 3806
 FT REPEAT 3810
 FT REPEAT 3814
 FT REPEAT 3818
 FT REPEAT 3822
 FT REPEAT 3826
 FT REPEAT 3830
 FT REPEAT 3834
 FT REPEAT 3838
 FT REPEAT 3842
 FT REPEAT 3846
 FT REPEAT 3850
 FT REPEAT 3854
 FT REPEAT 3858
 FT REPEAT 3862
 FT REPEAT 3866
 FT REPEAT 3870
 FT REPEAT 3874
 FT REPEAT 3878
 FT REPEAT 3882
 FT REPEAT 3886
 FT REPEAT 3890
 FT REPEAT 3894
 FT REPEAT 3898
 FT REPEAT 3902
 FT REPEAT 3906
 FT REPEAT 3910
 FT REPEAT 3914
 FT REPEAT 3918
 FT REPEAT 3922
 FT REPEAT 3926
 FT REPEAT 3930
 FT REPEAT 3934
 FT REPEAT 3938
 FT REPEAT 3942
 FT REPEAT 3946
 FT REPEAT 3950
 FT REPEAT 3954
 FT REPEAT 3958
 FT REPEAT 3962
 FT REPEAT 3966
 FT REPEAT 3970
 FT REPEAT 3974
 FT REPEAT 3978
 FT REPEAT 3982
 FT REPEAT 3986
 FT REPEAT 3990
 FT REPEAT 3994
 FT REPEAT 3998
 FT REPEAT 4002
 FT REPEAT 4006
 FT REPEAT 4010
 FT REPEAT 4014
 FT REPEAT 4018
 FT REPEAT 4022
 FT REPEAT 4026
 FT REPEAT 4030
 FT REPEAT 4034
 FT REPEAT 4038
 FT REPEAT 4042
 FT REPEAT 4046
 FT REPEAT 4050
 FT REPEAT 4054
 FT REPEAT 4058
 FT REPEAT 4062
 FT REPEAT 4066
 FT REPEAT 4070
 FT REPEAT 4074
 FT REPEAT 4078
 FT REPEAT 4082
 FT REPEAT 4086
 FT REPEAT 4090
 FT REPEAT 4094
 FT REPEAT 4098
 FT REPEAT 4102
 FT REPEAT 4106
 FT REPEAT 4110
 FT REPEAT 4114
 FT REPEAT 4118
 FT REPEAT 4122
 FT REPEAT 4126
 FT REPEAT 4130
 FT REPEAT 4134
 FT REPEAT 4138
 FT REPEAT 4142
 FT REPEAT 4146
 FT REPEAT 4150
 FT REPEAT 4154
 FT REPEAT 4158
 FT REPEAT 4162
 FT REPEAT 4166
 FT REPEAT 4170
 FT REPEAT 4174
 FT REPEAT 4178
 FT REPEAT 4182
 FT REPEAT 4186
 FT REPEAT 4190
 FT REPEAT 4194
 FT REPEAT 4198
 FT REPEAT 4202
 FT REPEAT 4206
 FT REPEAT 4210
 FT REPEAT 4214
 FT REPEAT 4218
 FT REPEAT 4222
 FT REPEAT 4226
 FT REPEAT 4230
 FT REPEAT 4234
 FT REPEAT 4238
 FT REPEAT 4242
 FT REPEAT 4246
 FT REPEAT 4250
 FT REPEAT 4254
 FT REPEAT 4258
 FT REPEAT 4262
 FT REPEAT 4266
 FT REPEAT 4270
 FT REPEAT 4274
 FT REPEAT 4278
 FT REPEAT 4282
 FT REPEAT 4286
 FT REPEAT 4290
 FT REPEAT 4294
 FT REPEAT 4298
 FT REPEAT 4302
 FT REPEAT 4306
 FT REPEAT 4310
 FT REPEAT 4314
 FT REPEAT 4318
 FT REPEAT 4322
 FT REPEAT 4326
 FT REPEAT 4330
 FT REPEAT 4334
 FT REPEAT 4338
 FT REPEAT 4342
 FT REPEAT 4346
 FT REPEAT 4350
 FT REPEAT 4354
 FT REPEAT 4358
 FT REPEAT 4362
 FT REPEAT 4366
 FT REPEAT 4370
 FT REPEAT 4374
 FT REPEAT 4378
 FT REPEAT 4382
 FT REPEAT 4386
 FT REPEAT 4390
 FT REPEAT 4394
 FT REPEAT 4398
 FT REPEAT 4402
 FT REPEAT 4406
 FT REPEAT 4410
 FT REPEAT 4414
 FT REPEAT 4418
 FT REPEAT 4422
 FT REPEAT 4426
 FT REPEAT 4430
 FT REPEAT 4434
 FT REPEAT 4438
 FT REPEAT 4442
 FT REPEAT 4446
 FT REPEAT 4450
 FT REPEAT 4454
 FT REPEAT 4458
 FT REPEAT 4462
 FT REPEAT 4466
 FT REPEAT 4470
 FT REPEAT 4474
 FT REPEAT 4478
 FT REPEAT 4482
 FT REPEAT 4486
 FT REPEAT 4490
 FT REPEAT 4494
 FT REPEAT 4498
 FT REPEAT 4502
 FT REPEAT 4506
 FT REPEAT 4510
 FT REPEAT 4514
 FT REPEAT 4518
 FT REPEAT 4522
 FT REPEAT 4526
 FT REPEAT 4530
 FT REPEAT 4534
 FT REPEAT 4538
 FT REPEAT 4542
 FT REPEAT 4546
 FT REPEAT 4550
 FT REPEAT 4554
 FT REPEAT 4558
 FT REPEAT 4562
 FT REPEAT 4566
 FT REPEAT 4570
 FT REPEAT 4574
 FT REPEAT 4578
 FT REPEAT 4582
 FT REPEAT 4586
 FT REPEAT 4590
 FT REPEAT 4594
 FT REPEAT 4598
 FT REPEAT 4602
 FT REPEAT 4606
 FT REPEAT 4610
 FT REPEAT 4614
 FT REPEAT 4618
 FT REPEAT 4622
 FT REPEAT 4626
 FT REPEAT 4630
 FT REPEAT 4634
 FT REPEAT 4638
 FT REPEAT 4642
 FT REPEAT 4646
 FT REPEAT 4650
 FT REPEAT 4654
 FT REPEAT 4658
 FT REPEAT 4662
 FT REPEAT 4666
 FT REPEAT 4670
 FT REPEAT 4674
 FT REPEAT 4678
 FT REPEAT 4682
 FT REPEAT 4686
 FT REPEAT 4690
 FT REPEAT 4694
 FT REPEAT 4698
 FT REPEAT 4702
 FT REPEAT 4706
 FT REPEAT 4710
 FT REPEAT 4714
 FT REPEAT 4718
 FT REPEAT 4722
 FT REPEAT 4726
 FT REPEAT 4730
 FT REPEAT 4734
 FT REPEAT 4738
 FT REPEAT 4742
 FT REPEAT 4746
 FT REPEAT 4750
 FT REPEAT 4754
 FT REPEAT 4758
 FT REPEAT 4762
 FT REPEAT 4766
 FT REPEAT 4770
 FT REPEAT 4774
 FT REPEAT 4778
 FT REPEAT 4782
 FT REPEAT 4786
 FT REPEAT 4790
 FT REPEAT 4794
 FT REPEAT 4798
 FT REPEAT 4802
 FT REPEAT 4806
 FT REPEAT 4810
 FT REPEAT 4814
 FT REPEAT 4818
 FT REPEAT 4822
 FT REPEAT 4826
 FT REPEAT 4830
 FT REPEAT 4834
 FT REPEAT 4838
 FT REPEAT 4842
 FT REPEAT 4846
 FT REPEAT 4850
 FT REPEAT 4854
 FT REPEAT 4858
 FT REPEAT 4862
 FT REPEAT 4866
 FT REPEAT 4870
 FT REPEAT 4874
 FT REPEAT 4878
 FT REPEAT 4882
 FT REPEAT 4886
 FT REPEAT 4890
 FT REPEAT 4894
 FT REPEAT 4898
 FT REPEAT 4902
 FT REPEAT 4906
 FT REPEAT 4910
 FT REPEAT 4914
 FT REPEAT 4918
 FT REPEAT 4922
 FT REPEAT 4926
 FT REPEAT 4930
 FT REPEAT 4934
 FT REPEAT 4938
 FT REPEAT 4942
 FT REPEAT 4946
 FT REPEAT 4950
 FT REPEAT 4954
 FT REPEAT 4958
 FT REPEAT 4962
 FT REPEAT 4966
 FT REPEAT 4970
 FT REPEAT 4974
 FT REPEAT 4978
 FT REPEAT 4982
 FT REPEAT 4986
 FT REPEAT 4990
 FT REPEAT 4994
 FT REPEAT 4998
 FT REPEAT 5002
 FT REPEAT 5006
 FT REPEAT 5010
 FT REPEAT 5014
 FT REPEAT 5018
 FT REPEAT 5022
 FT REPEAT 5026
 FT REPEAT 5030
 FT REPEAT 5034
 FT REPEAT 5038
 FT REPEAT 5042
 FT REPEAT 5046
 FT REPEAT 5050
 FT REPEAT 5054
 FT REPEAT 5058
 FT REPEAT 5062
 FT REPEAT 5066
 FT REPEAT 5070
 FT REPEAT 5074
 FT RE

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z68198; CAA92385.1; -
DR PROSITE: PS50077; HEAT_REPEAT: 4.
KW Hypothetical protein; Translation regulation; Activator; Repeat.
FT REPEAT 315 352 HEAT 1.
FT REPEAT 1062 1099 HEAT 2.
FT REPEAT 1319 1356 HEAT 3.
FT REPEAT 1439 1476 HEAT 4.
FT REPEAT 1478 1514 HEAT 5.
FT REPEAT 1518 1555 HEAT 6.
FT REPEAT 1557 1593 HEAT 7.
FT REPEAT 1637 1674 HEAT 8.
FT REPEAT 1676 1713 HEAT 9.
FT REPEAT 1714 1751 HEAT 10.
FT REPEAT 1755 1792 HEAT 11.
FT REPEAT 1793 1830 HEAT 12.
FT REPEAT 1898 1939 HEAT 13.
FT REPEAT 1941 1977 HEAT 14.
FT REPEAT 1982 2019 HEAT 15.
FT REPEAT 2020 2055 HEAT 16.
FT REPEAT 2057 2090 HEAT 17.
FT REPEAT 2319 2361 HEAT 18.
FT REPEAT 2379 2416 HEAT 19.
SQ SEQUENCE 2670 AA; 297333 MW; 948E9316D56D74C3 CRC64;

Query Match 1.7%; Score 8; DB 1; Length 2670;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 38 LSGLAADS 45
Db 1765 LSGLAADS 1772
|||||||

RESULT 23
SH_RAT ID SH_RAT STANDARD; PRT: 106 AA.
AC P55248;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE PUTATIVE PROTEIN SH (SH-4).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RT TISSUE=Heart;
RX MEDLINE=87149087; PubMed=3547652;
RA Adelman J.P., Bond C.T., Douglass J., Herbert E.;
RT "Two mammalian genes transcribed from opposite strands of the same
RT DNA locus".
RL Science 235:1514-1517(1987).
CC -!- FUNCTION: MAY BE INVOLVED WITH THE REGULATION OF GNRH GENE
CC EXPRESSION. IT IS NOT KNOWN IF THIS PROTEIN IS TRANSCRIBED.
CC -!- TISSUE SPECIFICITY: HEART.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

FT DISULFID 681 BY SIMILARITY.
FT DISULFID 686 BY SIMILARITY.
FT DISULFID 703 BY SIMILARITY.
FT DISULFID 719 BY SIMILARITY.
FT DISULFID 724 BY SIMILARITY.
FT DISULFID 740 BY SIMILARITY.
FT DISULFID 756 BY SIMILARITY.
FT DISULFID 761 BY SIMILARITY.
FT DISULFID 778 BY SIMILARITY.
FT DISULFID 794 BY SIMILARITY.
FT DISULFID 799 BY SIMILARITY.
FT DISULFID 816 BY SIMILARITY.
FT DISULFID 832 BY SIMILARITY.
FT DISULFID 837 BY SIMILARITY.
FT DISULFID 856 BY SIMILARITY.
FT DISULFID 872 BY SIMILARITY.
FT DISULFID 877 BY SIMILARITY.
FT DISULFID 894 BY SIMILARITY.
FT DISULFID 910 BY SIMILARITY.
FT DISULFID 915 BY SIMILARITY.
FT DISULFID 932 BY SIMILARITY.
FT DISULFID 986 BY SIMILARITY.
FT DISULFID 991 BY SIMILARITY.
FT DISULFID 1008 BY SIMILARITY.
FT DISULFID 1024 BY SIMILARITY.
FT DISULFID 1029 BY SIMILARITY.
FT DISULFID 1046 BY SIMILARITY.
FT DISULFID 1062 BY SIMILARITY.
FT DISULFID 1067 BY SIMILARITY.
FT DISULFID 1084 BY SIMILARITY.
FT DISULFID 1100 BY SIMILARITY.
FT DISULFID 1115 BY SIMILARITY.
FT DISULFID 1132 BY SIMILARITY.
FT DISULFID 1148 BY SIMILARITY.
FT DISULFID 1153 BY SIMILARITY.
FT DISULFID 1170 BY SIMILARITY.
FT DISULFID 1186 BY SIMILARITY.
FT DISULFID 1191 BY SIMILARITY.
FT DISULFID 1206 BY SIMILARITY.
FT DISULFID 1208 BY SIMILARITY.
FT DISULFID 1224 BY SIMILARITY.
FT DISULFID 1237 BY SIMILARITY.

Query Match 1.7%; Score 8; DB 1; Length 2524;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 PNPCHNGG 87
Db 912 PNPCHNGG 919
|||||||

RESULT 22
YAO5_SCHPO ID YAO5_SCHPO STANDARD; PRT: 2670 AA.
AC Q10105;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE PUTATIVE TRANSLATIONAL ACTIVATOR C1866.05C (GCN1 HOMOLOG).
GN SPAC1866.05C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RT STRAIN=972;
RA Connor R., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: STRONG, TO YEAST GCN1.
CC -!- SIMILARITY: CONTAINS 19 HEAT REPEATS.
CC -----

```
CC -----
DR EMBL; M15527; AAA42140.1;
SQ SEQUENCE 106 AA; 11792 MW; F9EF8FDBFEAABE3E CRC64;

Query Match 1.5%; Score 7; DB 1; Length 106;
Best Local Similarity 100.0%; Pred. No. 7.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 NGGICLS 39
   |||||
DB 100 NGGICLS 106

RESULT 24
CYOD_PSEPU STANDARD; PRT; 110 AA.
AC Q9WR4;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE CYTOCHROME O UBIQUINOL OXIDASE PROTEIN CYOD.
GN CYOD.
OS Pseudomonas putida.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IH-2000;
RX MEDLINE=99085656; PubMed=9868765;
RA Hirayama H., Takami H., Inoue A., Horikoshi K.;
RT "Isolation and characterization of toluene-sensitive mutants from
   Pseudomonas putida IH-2000."
RL FEMS Microbiol. Lett. 169:219-225(1998).
CC -!- FUNCTION: CYTOCHROME O TERMINAL OXIDASE COMPLEX IS THE COMPONENT
CC OF THE AEROBIC RESPIRATORY CHAIN THAT PREDOMINATES WHEN CELLS ARE
CC GROWN AT HIGH AERATION (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE
CC (BY SIMILARITY).
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL; AB016787; BAA76359.1;
KW Oxidoreductase; Electron transport; Transmembrane; Inner membrane.
FT DOMAIN 1 18 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 19 39 POTENTIAL.
FT DOMAIN 40 46 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 47 67 POTENTIAL.
FT DOMAIN 68 78 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 79 99 POTENTIAL.
FT DOMAIN 100 110 PERIPLASMIC (POTENTIAL).
SQ SEQUENCE 110 AA; 12376 MW; 074ED0BE854FC0C9 CRC64;

Query Match 1.5%; Score 7; DB 1; Length 110;
Best Local Similarity 100.0%; Pred. No. 8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 LLVGLSL 15
   |||||
DB 91 LLVGLSL 97

RESULT 25
CRAA_RANTE STANDARD; PRT; 149 AA.
ID CRAA_RANTE
```

```
AC P02508;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE ALPHA CRYSTALLIN A CHAIN (FRAGMENT).
GN CRYAA.
OS Rana temporaria (European common frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidea; Ranidae; Rana.
OX NCBI_TaxID=8407;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84005173; PubMed=6604666;
RA Tomarev S.I., Zinovieva R.D., Dolgilevich S.M., Krayev A.S.,
RA Skryabin K.G., Gause G.G. Jr.;
RT "The absence of the long 3'-non-translated region in mRNA coding for
   eye lens alpha A2-crystallin of the frog (Rana temporaria).";
RL FEBS Lett. 162:47-51(1983).
CC -!- FUNCTION: MAY CONTRIBUTE TO THE TRANSPARENCY AND REFRACTIVE INDEX
CC OF THE LENS.
CC -!- SIMILARITY: BELONGS TO THE SMALL HEAT SHOCK PROTEIN (HSP20)
CC FAMILY. STRONG TO ALPHA(B)-CRYSTALLIN.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL; X00716; CAA25308.1;
DR PIR; A02911; CYEGA2.
DR InterPro; IPR002068;
DR Fram; PF00011; HSP20; 1.
DR PROSITE; PS01031; HSP20; 1.
KW Eye lens protein; Glycoprotein.
FT NON_TER 1
FT CARBOHYD 138 138 O-LINKED (GLCNAC) (BY SIMILARITY).
SQ SEQUENCE 149 AA; 16949 MW; 877E89A7428DB4DC CRC64;

Query Match 1.5%; Score 7; DB 1; Length 149;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 69 EEKPTSA 75
   |||||
DB 140 EEKPTSA 146

Search completed: May 23, 2001, 06:25:03
Job time: 197 sec
```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 06:21:26 ; Search time 47.04 Seconds
(without alignments)
1195.999 Million cell updates/sec

Title: US-09-237-981-10
Perfect score: 480
Sequence: 1 MKHLVAALLVGLSLGVPQF.....WSWYGRITLRSLLGCAEE 480

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 374700 seqs, 117207915 residues

Word size : 0

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 50 summaries

Database : SPTREMBL15.*

- 1: sp-archaea.*
- 2: sp-bacteria.*
- 3: sp-fungi.*
- 4: sp-human.*
- 5: sp-invertebrate.*
- 6: sp-mammal.*
- 7: sp-mhc.*
- 8: sp-organelle.*
- 9: sp-phage.*
- 10: sp-plant.*
- 11: sp-rodent.*
- 12: sp-unclassified.*
- 13: sp-vertebrate.*
- 14: sp-virus.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	293	61.0	480	11	O35474
2	88	18.3	480	4	O43854
3	10	2.1	78	4	P78328
4	10	2.1	263	4	O99740
5	10	2.1	1193	13	Q90819
6	10	2.1	1218	4	O15122
7	10	2.1	1218	4	O15816
8	10	2.1	1218	4	O14902
9	10	2.1	1218	11	Q9QXX0
10	10	2.1	1219	11	Q63722
11	10	2.1	1227	4	P78504
12	9	1.9	363	6	O77718
13	9	1.9	426	11	Q9WFS3
14	9	1.9	463	11	Q9RLX9
15	9	1.9	779	4	Q9NTT3
16	9	1.9	858	5	O76470
17	9	1.9	921	11	Q9QX38
18	9	1.9	1404	5	Q9VB65
19	9	1.9	2146	5	Q9VC97

20	9	1.9	2343	6	O62730
21	9	1.9	2343	6	O18806
22	8	1.7	63	2	Q9RTW3
23	8	1.7	247	2	Q9RTW3
24	8	1.7	252	4	Q95378
25	8	1.7	330	4	Q9NT67
26	8	1.7	481	10	O9LYD5
27	8	1.7	529	5	Q25058
28	8	1.7	832	5	Q99108
29	8	1.7	1193	5	Q19617
30	8	1.7	1212	13	O42347
31	8	1.7	1242	4	Q9NS15
32	8	1.7	1253	11	O61810
33	8	1.7	1378	11	Q07314
34	8	1.7	1395	11	Q07313
35	8	1.7	1399	11	Q07280
36	8	1.7	1438	11	Q07312
37	8	1.7	1471	11	Q07311
38	8	1.7	1578	11	Q07310
39	8	1.7	1715	11	O63375
40	8	1.7	1728	11	O63374
41	8	1.7	1800	2	Q9L948
42	8	1.7	2183	11	O88783
43	8	1.7	2531	5	O16004
44	7	1.5	50	2	P77193
45	7	1.5	56	11	O9QWB5
46	7	1.5	96	5	O9XWN8
47	7	1.5	101	8	Q9ZTK5
48	7	1.5	107	2	Q9JQ07
49	7	1.5	125	2	Q9RAT2
50	7	1.5	135	2	O53977

ALIGNMENTS

RESULT	1
O35474	
ID	O35474 PRELIMINARY; PRT; 480 AA.
AC	O35474; O35475;
DT	01-JAN-1998 (TREMREL. 05, Created)
DT	01-JAN-1998 (TREMREL. 05, Last sequence update)
DT	01-OCT-2000 (TREMREL. 15, Last annotation update)
DE	INTEGRIN-BINDING PROTEIN DEL1 PRECURSOR.
GN	EDIL3 OR DEL1.
OS	Mus musculus (Mouse).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX	NCBI_TaxID=10090;
RN	[1]
RP	SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND CHARACTERIZATION.
RC	TISSUE=EMBRYO;
RX	MEDLINE=98083109; PubMed=9420328;
RA	Hidai C., Zupancic T., Penta K., Mikhail A., Kawana M.,
RA	Quertermous E.E., Aoka Y., Fukagawa M., Matsui Y., Platika D.,
RA	Auerbach R., Hogan B.L.M., Snodgrass R., Quertermous T.;
RT	"Cloning and characterization of developmental endothelial locus-1: an
RT	embryonic endothelial cell protein that binds the alphavbeta3 integrin
RT	receptor.";
RL	Genes Dev. 12:21-33(1998).
CC	-!- FUNCTION: PROMOTES ADHESION OF ENDOTHELIAL CELLS THROUGH
CC	INTERACTION WITH THE ALPHA-V BETA3 INTEGRIN RECEPTOR. INHIBITS
CC	FORMATION OF VASCULAR-LIKE STRUCTURES. MAY BE INVOLVED IN
CC	REGULATION OF VASCULAR MORPHOGENESIS OR REMODELING IN EMBRYONIC
CC	DEVELOPMENT.
CC	-!- SUBCELLULAR LOCATION: EXTRACELLULAR MATRIX.
CC	-!- ALTERNATIVE PRODUCTS: TWO ISOFORMS; LONG ISOFORM (SHOWN HERE) AND
CC	SHORT ISOFORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC	-!- TISSUE SPECIFICITY: EXPRESSED IN ANGIOBLASTS AND EARLY ENDOTHELIAL
CC	CELLS. BY EMBRYONIC DAY 13.5, ALSO EXPRESSED IN A RESTRICTED GROUP
CC	OF NON-ENDOTHELIAL CELLS INCLUDING CHONDROCYTES AND RETINAL
CC	NEURONS.
CC	-!- DEVELOPMENTAL STAGE: EXPRESSED IN THE EMBRYO FROM DAY 7. AFTER DAY

AC	ISSUE-SPIRAL CORO;	
RX	MEDLINE-96I75595; PubMed-8631496;	
RA	Myat A., Henrique D., Ish-Horowicz D., Lewis J.;	
RT	"A chick homologue of Serrate and its relationship with Notch and	
RT	Delta homologues during central neurogenesis.;"	
RL	Dev. Biol. 174:233-247(1996).	
DR	EMBL; X95283; CAA64604.1; -;	
DR	HSSP; P00740; 1EDM.	
DR	INTERPRO; IPR000083; -;	
DR	INTERPRO; IPR000152; -;	
DR	INTERPRO; IPR000561; -;	
DR	INTERPRO; IPR000742; -;	
DR	INTERPRO; IPR001010; -;	
DR	INTERPRO; IPR001438; -;	
DR	INTERPRO; IPR001774; -;	
DR	INTERPRO; IPR001891; -;	
DR	INTERPRO; IPR001947; -;	
DR	INTERPRO; IPR002049; -;	

DR	PFAM; PF000008; EGF; 14.
DR	PFAM; PF01414; DSL; 1.
DR	PRINTS; PR00010; EGFBLOOD.
DR	PRINTS; PR00011; EGFLAMININ.
DR	PRINTS; PR00012; FNTYPEI.
DR	PRINTS; PR00286; CHARYBDTOXIN.
DR	PRINTS; PR00287; THIONIN.
DR	PROSITE; PS00010; ASX_HYDROXYL; 10.
DR	PROSITE; PS00022; EGF_1; UNKNOWN_16.
DR	PROSITE; PS01186; EGF_2; 12.
DR	PROSITE; PS01187; EGF_CA; 8.
KW	Glycoprotein; EGF-like domain.
FT	NON_TER 1
SQ	SEQUENCE 1193 AA; 131039 MW; 24E94D09EB987935 CRC64;

Query Match		2.1%; Score 10; DB 13; Length 1193;
Best Local Similarity 100.0%; Pred. No. 0.2;		
Matches	10; Conservative	0; Mismatches 0; Indels 0; Gaps

QY	80 PNPCNGGTC 89
Dd	724 PNPCHNGTGC 733

RESULT	6
O15122	PRELIMINARY; PRT; 1218 AA.
ID	O15122 AC O15122
OC	01-JAN-1998 (TrEMBLrel. 05, Created)
DT	01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT	01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
JAGGED1.	
GN	JAG1.
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX	NCBI_TaxID=9606;
[1]	
RN	SEQUENCE FROM N.A.
RA	Oda T., Elkhoun A.G., Pike B.L., Okajima K., Krantz I.D., Genin A.,
RA	Piccoli D.A., Meltzer P.S., Spinner N.B., Collins F.S.,
RA	Chandrasekharappa S.C.;
RL	Nat. Genet. 0:0-0(1997).
[2]	
RX	SEQUENCE FROM N.A.
RP	MEDLINE=97422615; PubMed=9268641;
RA	Oda T., Elkhoun A.G., Meltzer P.S., Chandrasekharappa S.C.;
RT	"Identification and cloning of the human homolog (JAG1) of the rat
RT	jagged1 gene from the Alagille syndrome critical region at 20p12.";
RL	Genomics 43:376-379(1997).
DR	EMBL; AF003837; AAC51731.1; -.
DR	HSSP; P00740; 11XA.
DR	INTERPRO; IPR000152; -.
DR	INTERPRO; IPR000561; -.
DR	INTERPRO; IPR000742; -.
DR	INTERPRO; IPR001438; -.
DR	INTERPRO; IPR001774; -.
DR	INTERPRO; IPR001881; -.
DR	PFAM; PF000008; EGF; 14.
DR	PFAM; PF01414; DSL; 1.
DR	PRINTS; PR00010; EGFBLOOD.
DR	PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_10.
DR	PROSITE; PS00022; EGF_1; UNKNOWN_16.
DR	PROSITE; PS01186; EGF_2; 12.
DR	PROSITE; PS01187; EGF_CA; 8.
KW	Glycoprotein; EGF-like domain.
SQ	SEQUENCE 1218 AA; 133958 MW; F34BE15FE265377C CRC64;

Query Match		2.1%; Score 10; DB 4; Length 1218;
Best Local Similarity 100.0%; Pred. No. 0.2;		
Matches	10; Conservative	0; Mismatches 0; Indels 0; Gaps

EX MEDLINE=95211842; PubMed=7697721;
RA Lindsell C.E., Shawber C.J., Boulter J., Weinmaster G.;
RT "Jagged: a mammalian ligand that activates Notch1.";
RL Cell 80:909-917(1995).
RN [2]
RP SEQUENCE FROM N.A.
RA Bash J., Zong W.-X., Gelinas C.;
RL Submitted (Oct-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF028593; AAB84053.1; -.
DR HSSP; P00740; IIXA.
DR INTERPRO; IPR000152; -.
DR INTERPRO; IPR000561; -.
DR INTERPRO; IPR000742; -.
DR INTERPRO; IPR001438; -.
DR INTERPRO; IPR001774; -.
DR INTERPRO; IPR001881; -.
DR PFAM; PF00008; EGF_14.
DR PFAM; PF01414; DSL; 1.
DR PRINTS; PR00010; EGFBL00D.
DR PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_10.
DR PROSITE; PS00022; EGF_1; UNKNOWN_16.
DR PROSITE; PS01186; EGF_2; 12.
DR PROSITE; PS01187; EGF_CA; 8.
KW Transmembrane; Glycoprotein; EGF-like domain.
SQ SEQUENCE 1218 AA; 133797 MW; F36EE9FBF64DF162 CRC64;

Query Match 2.1%; Score 10; DB 4; Length 1218;
Best Local Similarity 100.0%; Pred. No. 0.2;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 PNPCHNGGTC 89
|||||
DB 750 PNPCHNGGTC 759

RESULT 9
ID Q9QXX0 PRELIMINARY; PRT; 1218 AA.
AC Q9QXX0:
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
DE JAGGED1.
GN JAG1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SWISS WEBSTER/NIH;
RA Shimizu K., Chiba S., Kumano K., Hosoya N., Takahashi T., Kanda Y.,
RA Hamada Y., Yazaki Y., Hirai H.;
RT "Mouse Jagged1 physically interacts with Notch2 and other Notch
RT receptors: assessment by quantitative methods.";
RL J. Biol. Chem. 274:32961-32969(1999).
DR EMBL; AF171092; AAF15505.1; -.
DR HSSP; P00740; IIXA.
DR MGD; MGI:1095416; Jag1.
DR INTERPRO; IPR000152; -.
DR INTERPRO; IPR000561; -.
DR INTERPRO; IPR000742; -.
DR INTERPRO; IPR001438; -.
DR INTERPRO; IPR001774; -.
DR INTERPRO; IPR001881; -.
DR INTERPRO; IPR002049; -.
DR PFAM; PF00008; EGF; 14.
DR PFAM; PF01414; DSL; 1.
DR PRINTS; PR00010; EGFBL00D.
DR PRINTS; PR00011; EGFBL00N.
DR PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_10.
DR PROSITE; PS00022; EGF_1; UNKNOWN_16.

DR PROSITE; PS01186; EGF_2; 12.
DR PROSITE; PS01187; EGF_CA; 8.
SQ SEQUENCE 1218 AA; 134163 MW; 77739F8928BB793C CRC64;

Query Match 2.1%; Score 10; DB 11; Length 1218;
Best Local Similarity 100.0%; Pred. No. 0.2;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 PNPCHNGGTC 89
|||||
DB 750 PNPCHNGGTC 759

RESULT 10
Q63722
ID Q63722 PRELIMINARY; PRT; 1219 AA.
AC Q63722; P70640;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-FEB-1997 (TReMBLrel. 02, Last sequence update)
DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)
DE JAGGED PROTEIN.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-SCIATIC NERVE;
RX MEDLINE=95211842; PubMed=7697721;
RA Lindsell C.E., Shawber C.J., Boulter J., Weinmaster G.;
RT "Jagged: a mammalian ligand that activates Notch1.";
RL Cell 80:909-917(1995).
DR EMBL; L38483; AAB06509.1; -.
DR HSSP; P00740; IIXA.
DR INTERPRO; IPR000152; -.
DR INTERPRO; IPR000561; -.
DR INTERPRO; IPR000742; -.
DR INTERPRO; IPR001438; -.
DR INTERPRO; IPR001774; -.
DR INTERPRO; IPR001881; -.
DR PFAM; PF00008; EGF; 14.
DR PFAM; PF01414; DSL; 1.
DR PRINTS; PR00010; EGFBL00D.
DR PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_10.
DR PROSITE; PS00022; EGF_1; UNKNOWN_16.
DR PROSITE; PS01186; EGF_2; 12.
DR PROSITE; PS01187; EGF_CA; 8.
KW Glycoprotein; EGF-like domain.
SQ SEQUENCE 1219 AA; 134325 MW; 65D4CFC238A0E204 CRC64;

Query Match 2.1%; Score 10; DB 11; Length 1219;
Best Local Similarity 100.0%; Pred. No. 0.2;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 PNPCHNGGTC 89
|||||
DB 750 PNPCHNGGTC 759

RESULT 11
P78504
ID P78504 PRELIMINARY; PRT; 1227 AA.
AC P78504:
DT 01-MAY-1997 (TReMBLrel. 03, Created)
DT 01-MAY-1997 (TReMBLrel. 03, Last sequence update)
DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
DE JAGGED 1 (TRANSMEMBRANE PROTEIN JAGGED).
GN HJ1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE MILK FAT GLOBULE GLYCOPROTEIN MFG-E8 LONG FORM.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BALB/C; TISSUE-MAMMARY GLAND;
 RA Oshima K., Aoki N., Negi M., Kishi M., Kitajima K., Matsuda T.;
 .RT "Lactation-dependent expression of an mRNA splice variant with an exon
 RT for a multiply O-glycosylated domain of mouse milk fat globule
 RT glycoprotein MFG-E8"; Commun. 254:522-528(1999).
 RL Biochem. Biophys. Res. Commun. 254:522-528(1999).
 DR EMBL; AB021130; BAA35180.1; -;
 DR HSSP; P00740; 11XA.
 DR INTERPRO; IPR000421; -;
 DR INTERPRO; IPR000561; -;
 DR INTERPRO; IPR001092; -;
 DR INTERPRO; IPR001438; -;
 DR PFAM; PF00008; EGF; 2.
 DR PFAM; PF00754; F5_F8_type_C; 2.
 DR PRINTS; PR00010; EGFBL00.
 DR PROSITE; PS00022; EGF_1; UNKNOWN_2.
 DR PROSITE; PS01186; EGF_2; 2.
 DR PROSITE; PS01285; FA58C_1; 2.
 DR PROSITE; PS01286; FA58C_2; 2.
 SQ SEQUENCE 463 AA; 51269 MW; D719D2BE09E6427 CRC64;

Query Match 1.9%; Score 9; DB 11; Length 463;
 Best Local Similarity 100.0%; Pred. No. 0.95; Indels 0; Gaps 0;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 388 VTGIITQGA 396
 Db 377 VTGIITQGA 385
 RESULT 15
 Q9NTT3 PRELIMINARY; PRT; 779 AA.
 AC Q9NTT3;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DE BA342D11.1 (NEUROFILIN 1) (FRAGMENT).
 GN NRPI.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Bird C.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL121748; CAB87954.1; -;
 FT NON_TER
 SQ SEQUENCE 779 AA; 86870 MW; 0F6DDE605A0BAB7A CRC64;

Query Match 1.9%; Score 9; DB 4; Length 779;
 Best Local Similarity 100.0%; Pred. No. 1.5;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 307 LRMELLCGE 315
 Db 432 LRMELLCGE 440

RESULT 16
 076470

ID 076470 PRELIMINARY; PRT; 858 AA.
 AC 076470;
 DT 01-NOV-1998 (TREMBlrel. 08, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DE 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE ECHINONECTIN.
 OS Lytechinus variegatus (Sea urchin).
 OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
 OC Echinoidea; Euechinoidea; Echinacea; Temnopleuroidea; Toxopneustidae;
 OC Lytechinus.
 OX NCBI_TaxID=7654;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA "Echinonectin";
 RT Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 RL EMBL; AF070482; AAC32598.2; -;
 DR HSSP; P00451; 1CFG.
 DR INTERPRO; IPR000421; -;
 DR INTERPRO; IPR001092; -;
 DR PFAM; PF00754; F5_F8_type_C; 4.
 DR PROSITE; PS01285; FA58C_1; 4.
 SQ SEQUENCE 858 AA; 94845 MW; 56EC43AB727ED848 CRC64;

Query Match 1.9%; Score 9; DB 5; Length 858;
 Best Local Similarity 100.0%; Pred. No. 1.7;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 225 RVTGVITQG 233
 Db 603 RVTGVITQG 611

RESULT 17
 Q9QX38 PRELIMINARY; PRT; 921 AA.
 ID Q9QX38;
 AC Q9QX38;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
 DE NEUROFILIN.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA He Z., Tessier-Lavigne M.;
 RT "Neuropilin is a receptor for the axonal chemorepellent semaphorin
 RT III.";
 RL Cell 90:739-751(1997).
 DR EMBL; AF018957; AAC53345.1; -;
 DR INTERPRO; IPR000421; -;
 DR INTERPRO; IPR000859; -;
 DR INTERPRO; IPR000998; -;
 DR INTERPRO; IPR001092; -;
 DR PFAM; PF00431; CUB; 2.
 DR PFAM; PF00625; NAM_1.
 DR PFAM; PF00754; F5_F8_type_C; 2.
 DR PRINTS; PR00020; MAMDOMAIN.
 DR PROSITE; PS00740; NAM_1; 1.
 DR PROSITE; PS01285; FA58C_1; 2.
 DR PROSITE; PS01286; FA58C_2; 2.
 DR PROSITE; PS50060; NAM_2; 1.
 SQ SEQUENCE 921 AA; 103052 MW; 58B29A9AA4978971 CRC64;

Query Match 1.9%; Score 9; DB 11; Length 921;
 Best Local Similarity 100.0%; Pred. No. 1.8;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 307 LRMELLCGE 315

```

DB 575 LRMELGCE 583
|||||||
RESULT 18
Q9VB65 Q9VB65 PRELIMINARY; PRT; 1404 AA.
ID Q9VB65;
AC Q9VB65;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE SER PROTEIN.
GN SER.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN-BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Hostin D., Houston K.A., Howland T.J., Hernandez J.R., Houck J.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spler E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RA "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
DR EMBL: AE003759; AAF56678.1;
DR HSSP; P00743; 1CCF.
DR FLYBASE: FBgn0004197; Ser.
DR INTERPRO: IPR000152;
DR INTERPRO: IPR000255;
DR INTERPRO: IPR000515;
DR INTERPRO: IPR000561;
DR INTERPRO: IPR000742;
DR INTERPRO: IPR001438;
DR INTERPRO: IPR001774;
DR INTERPRO: IPR001801;
DR INTERPRO: IPR002049;
DR PFAM: PF000008; EGF; 11.

```

```

DR PFAM: PF01414; DSL; 1.
DR PRINTS: PRO0010; EGFBL00D.
DR PRINTS: PRO0011; EGFAMININ.
DR PROSITE: PS00010; ASX_HYDROXYL; 7.
DR PROSITE: PS00012; PHOSPHOPANTHETHEINE; UNKNOWN_1.
DR PROSITE: PS00022; EGF_1; 14.
DR PROSITE: PS00402; BPD_TRANSP_INN_MEMBR; UNKNOWN_1.
DR PROSITE: PS01186; EGF_2; 8.
DR PROSITE: PS01187; EGF_CA; 5.
SQ SEQUENCE 1404 AA; 150342 MW; E988604001DAAC84 CRC64;

Query Match 1.9%; Score 9; DB 5; Length 1404;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 82 PCHNGGTCE 90
|||||||
DB 883 PCHNGGTCE 891

RESULT 19
Q9VC97 Q9VC97 PRELIMINARY; PRT; 2146 AA.
ID Q9VC97;
AC Q9VC97;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE CRB PROTEIN.
GN CRB.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN-BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Hostin D., Houston K.A., Howland T.J., Hernandez J.R., Houck J.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spler E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RA "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
DR EMBL: AE003759; AAF56678.1;
DR HSSP; P00743; 1CCF.
DR FLYBASE: FBgn0004197; Ser.
DR INTERPRO: IPR000152;
DR INTERPRO: IPR000255;
DR INTERPRO: IPR000515;
DR INTERPRO: IPR000561;
DR INTERPRO: IPR000742;
DR INTERPRO: IPR001438;
DR INTERPRO: IPR001774;
DR INTERPRO: IPR001801;
DR INTERPRO: IPR002049;
DR PFAM: PF000008; EGF; 11.

```

RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 DR EMBL; AE003747; AAF56276.1; -.
 DR HSSP; P00740; 11XA.
 DR FLYBASE; FBgn000368; crb.
 DR INTERPRO; IPR000152; -.
 DR INTERPRO; IPR000561; -.
 DR INTERPRO; IPR000742; -.
 DR INTERPRO; IPR001438; -.
 DR INTERPRO; IPR001791; -.
 DR INTERPRO; IPR001881; -.
 DR INTERPRO; IPR002049; -.
 DR PFAM; PF00008; EGF_27.
 DR PFAM; PF00054; laminin_G; 3.
 DR PRINTS; PR00010; EGFBL00D.
 DR PRINTS; PR00011; EGFLAMININ.
 DR PROSITE; PS00010; ASX_HYDROXYL; 16.
 DR PROSITE; PS00022; EGF_1; 25.
 DR PROSITE; PS01186; EGF_2; 17.
 DR PROSITE; PS01187; EGF_CA; 13.
 SQ SEQUENCE 2146 AA; 233570 MW; 8E23B9E32B761115 CRC64;

Query Match 1.9%; Score 9; DB 5; Length 2146;

Best Local Similarity 100.0%; Pred. No. 3.8; Mismatches 0; Indels 0; Gaps 0;

QY 48 CRCPEGFAG 56
 |||||

DB 2056 CECPEGFAG 2064

RESULT 20
 O62730 PRELIMINARY; PRT; 2343 AA.
 AC O62730;
 DT 01-AUG-1998 (TREMBLrel. 07, Created)
 DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
 DE FACTOR VIII.
 OS *Canis familiaris* (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 OX NCBI_TaxID=9615;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=KIDNEY, SPLEEN;
 RA Gordy P.W., Bowen R.A.;
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF049489; AAC05384.1; -.
 DR HSSP; P00451; 1CFG.
 DR INTERPRO; IPR000421; -.
 DR INTERPRO; IPR001117; -.
 DR PFAM; PF00394; Cu-oxidase; 3.
 DR PFAM; PF00754; F5_F8_type_C; 2.
 DR PROSITE; PS00079; MULTICOPPER_OXIDASE1; 3.
 DR PROSITE; PS01285; FA58C.1; 2.
 DR PROSITE; PS01286; FA58C.2; 2.
 SQ SEQUENCE 2343 AA; 265613 MW; F612D744ADAADD99 CRC64;

Query Match 1.9%; Score 9; DB 6; Length 2343;

Best Local Similarity 100.0%; Pred. No. 4.1; Mismatches 0; Indels 0; Gaps 0;

QY 306 TLRMELGCG 314
 |||||

DB 2172 TLRMELGCG 2180

RESULT 21

OL18806 PRELIMINARY; PRT; 2343 AA.
 AC OL18806;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
 DE FACTOR VIII.
 DN F8.
 OS *Canis familiaris* (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 OX NCBI_TaxID=9615;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RA Cameron C., Notley C., Hoyle S., McGlynn L., Hough C., Kamisue S.,
 RA Giles A., Lillcrap D.;
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF016234; AAB87412.1; -.
 DR HSSP; P00451; 1CFG.
 DR INTERPRO; IPR000421; -.
 DR INTERPRO; IPR001117; -.
 DR PFAM; PF00394; Cu-oxidase; 3.
 DR PFAM; PF00754; F5_F8_type_C; 2.
 DR PROSITE; PS00079; MULTICOPPER_OXIDASE1; 3.
 DR PROSITE; PS01285; FA58C.1; 2.
 DR PROSITE; PS01286; FA58C.2; 2.
 SQ SEQUENCE 2343 AA; 265829 MW; A854FAE571C3B399 CRC64;

Query Match 1.9%; Score 9; DB 6; Length 2343;

Best Local Similarity 100.0%; Pred. No. 4.1; Mismatches 0; Indels 0; Gaps 0;

QY 306 TLRMELGCG 314
 |||||

DB 2172 TLRMELGCG 2180

RESULT 22
 Q9RTW3 PRELIMINARY; PRT; 63 AA.
 AC Q9RTW3;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
 DE HYPOTHETICAL 7.0 KDA PROTEIN.
 GN DRI640.
 OS *Deinococcus radiodurans*.
 OC Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.
 OX NCBI_TaxID=1299;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=RI;
 RX MEDLINE=20036896; PubMed=10567266;
 RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
 RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
 RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
 RA Vamathevan J.J., Lam P., McDonald L., Otterback T., Zaleski C.,
 RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
 RA Katchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
 RA Fraser C.M.;
 RT "Genome Sequence of the Radioresistant Bacterium *Deinococcus radiodurans* RI.";
 RL Science 286:1571-1577(1999).
 DR EMBL; AE002007; AAF11201.1; -.
 DR TIGR; DRI640; -.
 KW Hypothetical protein.
 SQ SEQUENCE 63 AA; 6982 MW; 116456AA28DD5BD CRC64;

Query Match 1.7%; Score 8; DB 2; Length 63;

Best Local Similarity 100.0%; Pred. No. 1.8;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 AWWLVGLS 14

Db 44 AWWLVGLS 51

RESULT 23

Q9R215 ID Q9R215 PRELIMINARY; PRT; 247 AA.
AC Q9R215;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE AMINO ACID ABC TRANSPORTER, PERMEASE PROTEIN.
GN DRA0138.
OS Deinococcus radiodurans.
OC Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RJ;
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zaleski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome Sequence of the Radioresistant Bacterium Deinococcus
RT radiodurans R1"; 1577(1999).
RL Science 286:1571-1577(1999).
DR EMBL; AE001862; AAF12222.1; -.
DR TIGR; DRA0138; -.
DR INTERPRO; IPR000515; -.
DR PFAM; PF00528; BPD_transp; 1.
SQ SEQUENCE 247 AA; 26293 MW; 80E694BE82938282 CRC64;

Query Match 1.7%; Score 8; DB 2; Length 247;

Best Local Similarity 100.0%; Pred. No. 6.1;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9 LLVGLSLG 16

Db 67 LLVGLSLG 74

RESULT 24

O95378 ID O95378 PRELIMINARY; PRT; 252 AA.
AC O95378;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE NEUREXIN III-ALPHA (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Rowen L., Madan A., Qin S., Ratcliffe A., Shaffer T., Abbasi N.,
RA Dickhoff R., James R., Loretz C., Madan A., Dors M., Hood L.;
RT "Sequence of chromosome 14 from 14q24.3-14q32.";
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF099810; AAC68909.1; -.
DR INTERPRO; IPR000901; -.
DR INTERPRO; IPR001791; -.
DR PFAM; PF00054; laminin_G; 1.
DR PROSITE; PS00442; GATASE_TYPE_I; UNKNOWN_1.
FT NON_TER 252

SQ SEQUENCE 252 AA; 27508 MW; F805A8A013C836D6 CRC64;

Query Match 1.7%; Score 8; DB 4; Length 252;
Best Local Similarity 100.0%; Pred. No. 6.2;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 30 PCENGGIC 37

Db 206 PCENGGIC 213

RESULT 25

Q9NT67 ID Q9NT67 PRELIMINARY; PRT; 330 AA.
AC Q9NT67;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE HYPOTHETICAL 36.5 KDA PROTEIN (FRAGMENT).
GN DKFZP761L191.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=AMYGDALA;
RA Ottenwaelder B., Obermaier B., Mewes H.W., Gassenhuber J., Wiemann S.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL137504; CAB70776.1; -.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 330 AA; 36465 MW; CAA0BF2DB9C7552D CRC64;

Query Match 1.7%; Score 8; DB 4; Length 330;

Best Local Similarity 100.0%; Pred. No. 7.9;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 82 PCHNGGTC 89

Db 73 PCHNGGTC 80

Search completed: May 23, 2001, 06:24:41

Job time: 195 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 06:14:01 ; Search time 35.38 seconds
(without alignments)
775.533 Million cell updates/sec

Title: US-09-237-981-10
Perfect score: 480
Sequence: 1 MKHLVAALLVGLSLGVPQF.....WSWYGRITRLSELLOCAEE 480

Scoring table: OLIGO

Gapop 60.0 ; Gapext 60.0

Searched: 390729 seqs, 57163235 residues

Word size : 0

Total number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 50 summaries

Database : A_Geneseq_0401.*

1: /SID56/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID56/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID56/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID56/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID56/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID56/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID56/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID56/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID56/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID56/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID56/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID56/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID56/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID56/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID56/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID56/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID56/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID56/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID56/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID56/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID56/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID56/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	480	100.0	480	W10364	Mouse developmenta
2	480	100.0	480	W94683	Murine Del-1 prote
3	217	45.2	221	W10366	Murine Del-1 trunc
4	194	40.4	221	W94684	Truncated murine D
5	79	16.5	513	W94685	Human Del-1 protei
6	57	11.9	85	W10367	Del-1 discoidin I-
7	57	11.9	85	W94689	Discoidin I/factor
8	54	11.2	481	W10365	Human developmenta
9	45	9.4	102	W10370	Human Del-1 splice
10	41	8.5	43	W94687	Del-1 epidermal gr
11	40	8.3	321	W94697	Human milk fat glo

12	18	3.8	25	18	W10368	Human Del-1 signal
13	18	3.8	25	20	W94698	Human Del-1 signal
14	18	3.8	42	20	W94688	Del-1 epidermal gr
15	14	2.9	57	20	W94686	Del-1 epidermal gr
16	10	2.1	1010	20	W87896	Human JAGGED1 solu
17	10	2.1	1036	18	W18351	Proliferation and
18	10	2.1	1187	18	W18352	Proliferation and
19	10	2.1	1193	17	W05835	Chick Serrate, Ga
20	10	2.1	1193	21	W59599	Human Jagged prote
21	10	2.1	1208	19	W40827	Human Jagged prote
22	10	2.1	1218	17	W05833	Human Serrate-1 (H
23	10	2.1	1218	18	W18354	Proliferation and
24	10	2.1	1218	19	W44301	Human Serrate 1.
25	10	2.1	1218	20	W87894	Human JAGGED1 prot
26	10	2.1	1218	21	W59597	Human Serrate prot
27	9	1.9	85	20	W94692	Discoidin I/factor
28	9	1.9	217	16	R77253	HMFG 46 kDa antige
29	9	1.9	218	16	R77254	HMFG 46 kDa antige
30	9	1.9	320	20	W98891	Peptide SEQ ID NO:
31	9	1.9	387	16	R77252	HMFG 46 kDa antige
32	9	1.9	387	21	Y94453	Human lactadherin
33	9	1.9	426	21	Y94454	Mouse lactadherin
34	9	1.9	644	20	Y06319	Human soluble neur
35	9	1.9	892	21	B24215	Soluble neuropilin
36	9	1.9	921	20	W96247	Rat semaphorin rec
37	9	1.9	922	20	W96309	Neuropilin. Rattu
38	9	1.9	923	20	Y06317	Human neuropilin-1
39	9	1.9	923	20	Y23247	Human VEGF165R/NP-
40	9	1.9	923	20	W96246	Human semaphorin r
41	9	1.9	923	20	W96248	Mouse semaphorin r
42	9	1.9	923	21	B24213	Mouse soluble neur
43	9	1.9	1404	14	R38304	Sequence of a serr
44	9	1.9	1404	21	Y59600	Drosophila Serrate
45	9	1.9	1530	21	B28151	Slit protein. Xen
46	9	1.9	2304	21	Y57848	Mouse factor VIII
47	9	1.9	2319	19	W53485	Murine factor VIII
48	9	1.9	2319	19	W44135	Mus musculus facto
49	9	1.9	2319	20	Y31596	Mouse factor VIII
50	9	1.9	2319	22	B50467	Mouse factor VIII.

ALIGNMENTS

RESULT 1
W10364
ID W10364 standard; protein; 480 AA.
XX
AC W10364;
XX
DT 03-MAY-1997 (first entry)
XX
DE Mouse developmentally-regulated endothelial cell locus-1 protein.
DE
KW Del-1; developmentally-regulated endothelial cell locus-1;
KW signal transduction; cancer; tumour marker; angiogenesis;
KW diabetic retinopathy; rheumatoid arthritis; endometriosis;
KW cardiac ischaemia; stroke; vascular disease; wound healing;
KW vulneryary; bone formation; diagnosis; therapy.
XX
OS Mus sp.
XX
FH Key Location/Qualifiers
FT Peptide 1..23
FT /label= Sig_peptide
FT Protein 24..480
FT /label= Mat_protein
FT Domain 26..59
FT /label= EGF-1
FT /note= "epidermal growth factor-like domain 1"
FT Domain 78..116
FT /label= EGF-2
FT /note= "epidermal growth factor-like domain 2"

```

FT Domain 123...154
FT /label= EGF-3
FT /note= "epidermal growth factor-like domain 3"
FT Domain 158...314
FT /label= Discoidin-1
FT /note= "discoidin I/factor VIII-like domain 1"
FT Domain 319...476
FT /label= Discoidin-2
FT /note= "discoidin I/factor VIII-like domain 2"
FT wO9640769-A1.
FT
FT 19-DEC-1996.
FT
FT 05-JUN-1996; 96WO-US09456.
FT
FT 07-JUN-1995; 95US-0480229.
FT
FT (PROG-) PROGENITOR INC.
FT (UYVA-) UNIV VANDERBILT.
FT
FT Hogan B, Quertermous T, Snodgrass HR, Zupancic TJ;
FT
FT WPI; 1997-052233/05.
FT N-PSDB; T47338.
FT
FT New developmentally regulated endothelial cell locus-1 (del-1) gene
FT - used to develop prods. for the diagnosis and treatment of cancer
FT and conditions involving abnormal angiogenesis
FT
FT Claim 3; Fig 6; 137pp; English.
FT
FT Murine Del-1 (W10364) is the polypeptide product of the murine
FT developmentally-regulated endothelial cell locus-1 (del-1) gene
FT (T47338). It shows 94% amino acid homology to the human Del-1
FT protein (W10365). Structurally, members of this novel gene family
FT contain 3 EGF-like domains and 2 discoidin I/factor VIII-like
FT domains. Del-1 is expressed in endothelial and certain tumour
FT cells. Its ability to inhibit vascular formation allows its used
FT as an anti-angiogenic agent. It can be used as a tumour marker,
FT to identify Del-1 binding partners, and to modulate endothelial
FT cell growth and blood vessel formation. Recombinant Del-1 can be
FT produced in transformed host cells utilising vectors incorporating
FT del-1 nucleic acids.
FT
FT Sequence 480 AA;
FT
FT
FT Query Match 100.0%; Score 480; DB 18; Length 480;
FT Best Local Similarity 100.0%; Pred. No. 0;
FT Matches 480; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
FT
QY 1 MKHLVAALLVGLSLGVDPFGKGDICNPNCENGICLSGLADSFSCCEPGEFAGPNC 60
DB 1 mkhlvaallvlgslgvdpfgkgdncnppncengiclsgladfsfsccepgfagpnc 60
QY 61 SVVEASDEKPTSAQCIPNPNCHNGGTCEISEAYRGDTFGYCKCPRGFNGTHCOHNI 120
DB 61 svveasdeekptsaqcipcnpnchnggtceiseayrgdtfgyckcprgfnghcqhni 120
QY 121 NECEAEPCRNNGICTDLVANYSCGPCGFMGRNCOYKCSGHLGTEGGIISNQITASSNH 180
DB 121 neceaeprnggictdlvanyscpcgfmgrncocykcsghlgteggliisnqgitassnh 180
QY 181 RALFLGLQKWPYYARLNKGLINAWTAENDRPWQIQLQKMRVTGVITOGAKRIGSP 240
DB 181 ralflglqkwpyyarlnkglinawtaendrpwqiqlqkmrvtgvtogakrigsp 240
QY 241 EYTKSYKIAYSDNGKTMWYKVGKNEEMVFRGNVDNNTPYANSFTTPPIKAQYVRLYPQI 300
DB 241 eytksykiaysdngktwmykvkgkneemvfrgnvndnntpyansfttpipikaqyvrlypqi 300
QY 301 CRRHCTLRMELLCGELSGCSEPLGMKSGHIDYQITASSVFTLNMDMFTWEPKRLDK 360

```

```

DB 301 crrhctlrnellgcslgscseplgmksghidytassvftlnmdmftwepkrlidk 360
QY 361 QGKYNAMWTSCHNDQSQWLQVLDLVPKVTGIIITQGAKEFGHVQFVGSYKLAYSNDGEHWM 420
DB 361 qgkynawtsghndsqswlqvdlvptkvtgiiitqgakdfghvqfvgyskylaysndgehw 420
QY 421 VHDEKORKKVKVQGNPDNTHRKNVIDPPIYARFTRILPWSWYGRITRLSELGCAEE 480
DB 421 vhdckgrkdkvfgnfdndthrkndvdpdiyarftrilpwsygritrlseligcaeee 480

RESULT 2
W94683
ID W94683 standard; Protein; 480 AA.
XX
AC W94683;
XX
DT 05-MAY-1999 (first entry)
XX
DE Murine Del-1 protein.
XX
KW Del-1; developmentally-regulated endothelial cell locus 1; cancer;
KW discoidin I; factor VIII-like domain; epidermal growth factor; EGF;
KW diabetic retinopathy; rheumatoid arthritis; endometriosis;
KW angiogenesis.
XX
OS Mus sp.
XX
PN US5877281-A.
XX
PD 02-MAR-1999.
XX
PF 05-JUN-1996; 96US-0659235.
XX
PR 05-JUN-1996; 96US-0659235.
XX
PR 07-JUN-1995; 95US-0480229.
XX
PA (PROG-) PROGENITOR INC.
XX (UYVA-) UNIV VANDERBILT.
XX
PI Hogan B, Quertermous T, Snodgrass HR, Zupancic TJ;
XX
DR WPI; 1999-189720/16.
DR N-PSDB; X18506.
XX
PT Del-1 polypeptide sequences - useful for treatment of cancer,
PT diabetic retinopathy, rheumatoid arthritis and endometriosis
XX
PS Claim 1; Fig 3; 73pp; English.
XX
CC The present sequence is murine developmentally-regulated endothelial
CC cell locus 1 (del-1). The protein has epidermal growth factor (EGF) like
CC domains and discoidin I/factor VIII-like domains. The del-1 proteins
CC have an inhibitory effect on angiogenesis (blood vessel growth), this
CC activity may be useful clinically to prevent neovascularisation of
CC tissues such as tumour nodules and prevention of metastases. The anti-
CC angiogenic activity of Del-1 may be used to treat abnormal conditions
CC that result from angiogenesis, including cancer, diabetic retinopathy,
CC rheumatoid arthritis and endometriosis. Since Del-1 promotes
CC angiogenesis it can be used to treat cardiac ischaemia, thrombotic
CC stroke, wound healing and peripheral vascular disease. Del-1 is also
CC useful for promoting bone formation. Del-1 binds to alpha V beta 3
CC integrin, and is an apoptosis inducer.
XX
SQ Sequence 480 AA;

```

```

Query Match 100.0%; Score 480; DB 20; Length 480;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 480; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MKHLVAALLVGLSLGVDPFGKGDICNPNCENGICLSGLADSFSCCEPGEFAGPNC 60

```



```

PR 05-JUN-1996; 960S-0659235.
PR 07-JUN-1995; 950S-0480229.
PA (PROG-) PROGENITOR INC.
PA (UYVA-) UNIV VANDERBILT.
XX
XX Hogan B, Quertermous T, Snodgrass HR, Zupancic TJ;
XX WPI; 1999-189720/16.
XX DR N-PSDB; X18507.
XX
XX Del-1 polypeptide sequences - useful for treatment of cancer,
XX PT diabetic retinopathy, rheumatoid arthritis and endometriosis
XX
XX Claim 3; Column 71-72; 73pp; English.
XX
XX The present sequence is truncated murine developmentally-regulated
XX CC endothelial cell locus 1 (Del-1). The protein has epidermal growth factor
XX CC (EGF) like domains and discoidin I/factor VIII-like domains. The Del-1
XX CC proteins have an inhibitory effect on angiogenesis (blood vessel growth),
XX CC this activity may be useful clinically to prevent neovascularisation of
XX CC tissues such as tumour nodules and prevention of metastases. The anti-
XX CC angiogenic activity of Del-1 may be used to treat abnormal conditions
XX CC that result from angiogenesis, including cancer, diabetic retinopathy,
XX CC rheumatoid arthritis and endometriosis. Since Del-1 promotes
XX CC angiogenesis it can be used to treat cardiac ischaemia, thrombotic
XX CC stroke, wound healing and peripheral vascular disease. Del-1 is also
XX CC useful for promoting bone formation. Del-1 binds to alpha v beta 3
XX CC Integrin, and is an apoptosis inducer.
XX
XX SQ Sequence 221 AA;
    Query Match 40.4%; Score 194; DB 20; Length 221;
    Best Local Similarity 100.0%; Pred. No. 1.1e-173;
    Matches 194; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKHLVAAILLVLSLGVPOFGKGDICNPENCGICSLGLADDSFSCPEGFAGPNC60
DB 1 mkhlvaawllvlgslgvpgdkgdncpncpgnggiclsgladdsfscpegfagpncs 60
QY 61 SVVEVASDEEKPSTAGPCIPNCPCHNGGTCTSEAYRGDTFGYVCKPCRGFNGHCHQNI 120
DB 61 svvevasdeekptsagpcipnchnggtctseayrgdtfgyvckpcrgfngihcqhni 120
QY 121 NECEAPCRNGGICTDLVANSCECPGFPMGRNCOYKCSGHLGIEGGIISNQOITASSNH 180
DB 121 neceapcrnggictdlvanscecpgfpmgrncqykcshlgieggiisnqqitassnh 180
QY 181 RALFLGLQKWYPYVA 194
DB 181 ralflglqkwypyya 194

RESULT 5
W94685
ID W94685 standard; Protein; 513 AA.
XX
XX AC W94685;
XX
XX 05-MAY-1999 (first entry)
XX
XX Ruman Del-1 protein.
XX
XX Del-1; developmentally-regulated endothelial cell locus 1; cancer;
XX KW discoidin I; factor VIII-like domain; epidermal growth factor; EGF;
XX KW diabetic retinopathy; rheumatoid arthritis; endometriosis;
XX KW angiogenesis.
XX
XX OS Homo sapiens.
XX
XX US5877281-A.
XX

```

```

PD 02-MAR-1999.
XX
XX 05-JUN-1996; 960S-0659235.
XX
XX 05-JUN-1996; 960S-0659235.
XX 07-JUN-1995; 950S-0480229.
XX
XX (PROG-) PROGENITOR INC.
XX PA (UYVA-) UNIV VANDERBILT.
XX
XX Hogan B, Quertermous T, Snodgrass HR, Zupancic TJ;
XX WPI; 1999-189720/16.
XX DR N-PSDB; X18508.
XX
XX Del-1 polypeptide sequences - useful for treatment of cancer,
XX PT diabetic retinopathy, rheumatoid arthritis and endometriosis
XX
XX Claim 5; Fig 4; 73pp; English.
XX
XX The present sequence is human developmentally-regulated endothelial
XX CC cell locus 1 (Del-1). The protein has epidermal growth factor (EGF) like
XX CC domains and discoidin I/factor VIII-like domains. The Del-1 proteins
XX CC have an inhibitory effect on angiogenesis (blood vessel growth), this
XX CC activity may be useful clinically to prevent neovascularisation of
XX CC tissues such as tumour nodules and prevention of metastases. The anti-
XX CC angiogenic activity of Del-1 may be used to treat abnormal conditions
XX CC that result from angiogenesis, including cancer, diabetic retinopathy,
XX CC rheumatoid arthritis and endometriosis. Since Del-1 promotes
XX CC angiogenesis it can be used to treat cardiac ischaemia, thrombotic
XX CC stroke, wound healing and peripheral vascular disease. Del-1 is also
XX CC useful for promoting bone formation. Del-1 binds to alpha v beta 3
XX CC Integrin, and is an apoptosis inducer.
XX
XX SQ Sequence 513 AA;
    Query Match 16.5%; Score 79; DB 20; Length 513;
    Best Local Similarity 100.0%; Pred. No. 1.1e-65;
    Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 341 FRTLNMDMTWEPKRLDKQKGVNAWTSCHNDQSQWLQVLLVPTKVTGIIQTQAKDFG 400
DB 374 frtlnmdmftwepkrldkqgkvnawtsgndqsqwlqvdllyptkvtglttgakdfg 433
QY 401 HVQFVGSYKLAISNDGEHW 419
DB 434 hvqfvgsyklaysndgehw 452

RESULT 6
W10367
ID W10367 standard; Protein; 85 AA.
XX
XX AC W10367;
XX
XX 03-MAY-1997 (first entry)
XX
XX Del-1 discoidin I-like domain.
XX
XX Del-1; developmentally-regulated endothelial cell locus-1;
XX KW signal transduction; cancer; tumour marker; angiogenesis;
XX KW diabetic retinopathy; rheumatoid arthritis; endometriosis;
XX KW cardiac ischaemia; stroke; vascular disease; wound healing;
XX KW vulnery; bone formation; diagnosis; therapy; discoidin.
XX
XX OS Mus sp.
XX
XX FH Key Location/Qualifiers
XX FT Misc-difference 16 /note= "unspecified amino acid residue"
XX FT Misc-difference 17 /note= "unspecified amino acid residue"
XX FT

```

FT Misc-difference 84
 FT PA /note= "unspecified amino acid residue"
 FT FT Misc-difference 85
 FT FT /note= "unspecified amino acid residue"
 XX W09640769-A1.
 XX PN
 XX PD 19-DEC-1996.
 XX PF 05-JUN-1996; 96WO-US09456.
 XX PR 07-JUN-1995; 95US-0480229.
 XX PA (PROG-) PROGENITOR INC.
 XX PA (UYVA-) UNIV VANDERBILT.
 XX PI Hogan B, Quertermous T, Snodgrass HR, Zupancic TJ;
 XX WPI; 1997-052233/05.
 XX DR New developmentally regulated endothelial cell locus-1 (del-1) gene
 PT - used to develop prods. for the diagnosis and treatment of cancer
 PT and conditions involving abnormal angiogenesis
 XX Example; Page 71-72; 137pp; English.
 XX CC A discoidin-like domain (W10367) was identified in mouse
 CC developmentally-regulated endothelial cell locus-1 (Del-1)
 CC sequences (see also W10364) derived from a trapped exon and
 CC mouse embryo cDNAs. It showed homology to discoidin-like domains
 CC from human milk fat globule protein, human factor V, mouse factor
 CC VIII, X-A5b1 and X-A5b2 (domains of Xenopus neuronal antigen A5)
 CC and discoidin I.
 XX CC
 XX SQ Sequence 85 AA;
 Query Match 11.9%; Score 57; DB 18; Length 85;
 Best Local Similarity 100.0%; Pred. No. 8.8e-46;
 Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 404 FVGSYKLAYSNDGHEWVHDEKQKDKVFGNFDNDTHRKNVIDPPIYARFIRILP 460
 Db 26 fvgysklaysndgehmvmvhdqekqrkdkvfgnfdndthrkndvdpdiyarfirilp 82
 RESULT 7
 W94689
 ID W94689 standard; peptide; 85 AA.
 XX AC W94689;
 XX DT 05-MAY-1999 (first entry)
 XX DE Discoidin I/factor VIII like domain #1.
 XX KW Del-1; developmentally-regulated endothelial cell locus 1; cancer;
 KW discoidin I; factor VIII-like domain; epidermal growth factor; EGF;
 KW diabetic retinopathy; rheumatoid arthritis; endometriosis;
 KW angiogenesis.
 XX OS Mus sp.
 XX OS Synthetic.
 XX PN US5877281-A.
 XX PD 02-MAR-1999.
 XX PF 05-JUN-1996; 96US-0659235.
 XX PR 05-JUN-1996; 96US-0659235.
 XX PR 07-JUN-1995; 95US-0480229.
 XX FT

PA (PROG-) PROGENITOR INC.
 PA (UYVA-) UNIV VANDERBILT.
 XX Hogan B, Quertermous T, Snodgrass HR, Zupancic TJ;
 XX WPI; 1999-189720/16.
 XX Del-1 polypeptide sequences - useful for treatment of cancer,
 PT diabetic retinopathy, rheumatoid arthritis and endometriosis
 XX Example; Column 37-38; 73pp; English.
 XX CC The present sequence represents a discoidin I/factor VIII-like domain.
 CC The present invention describes developmentally-regulated endothelial
 CC cell locus 1 (Del-1). Del-1 has epidermal growth factor like domains and
 CC discoidin I/factor VIII-like domains. The Del-1 proteins have an
 CC inhibitory effect on angiogenesis (blood vessel growth), this activity
 CC may be useful clinically to prevent neovascularisation of tissues such
 CC as tumour nodules and prevention of metastases. The anti-angiogenic
 CC activity of Del-1 may be used to treat abnormal conditions that result
 CC from angiogenesis, including cancer, diabetic retinopathy, rheumatoid
 CC arthritis and endometriosis. Since Del-1 promotes angiogenesis it can be
 CC used to treat cardiac ischaemia, thrombotic stroke, wound healing and
 CC peripheral vascular disease. Del-1 is also useful for promoting bone
 CC formation. Del-1 binds to alpha V beta 3 integrin, and is an apoptosis
 CC inducer.
 XX SQ Sequence 85 AA;
 Query Match 11.9%; Score 57; DB 20; Length 85;
 Best Local Similarity 100.0%; Pred. No. 8.8e-46;
 Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 404 FVGSYKLAYSNDGHEWVHDEKQKDKVFGNFDNDTHRKNVIDPPIYARFIRILP 460
 Db 26 fvgysklaysndgehmvmvhdqekqrkdkvfgnfdndthrkndvdpdiyarfirilp 82
 RESULT 8
 W10365
 ID W10365 standard; Protein; 481 AA.
 XX AC W10365;
 XX DT 03-MAY-1997 (first entry)
 XX DE Human developmentally-regulated endothelial cell locus-1 protein.
 XX KW Del-1; developmentally-regulated endothelial cell locus-1;
 KW signal transduction; cancer; tumour marker; angiogenesis;
 KW diabetic retinopathy; rheumatoid arthritis; endometriosis;
 KW cardiac ischaemia; stroke; vascular disease; wound healing;
 KW vulnery; bone formation; diagnosis; therapy.
 XX OS Homo sapiens.
 XX FH Key
 FH Peptide 3..23
 FT /label= Sig_peptide
 FT Protein 24..481
 FT /label= Mat_protein
 FT Domain 26..59
 FT /label= EGF-1
 FT /note= "epidermal growth factor-like domain 1"
 FT Domain 78..116
 FT /label= EGF-2
 FT /note= "epidermal growth factor-like domain 2"
 FT Domain 123..154
 FT /label= EGF-3
 FT /note= "epidermal growth factor-like domain 3"
 FT Domain 158..315
 FT /label= Discoidin-1

FT Domain /note= "discoidein I/factor VIII-like domain 1"
 FT 320..477
 FT /label= Discoidein-2
 FT /note= "discoidein I/factor VIII-like domain 2"
 XX
 PN W09640769-A1.
 XX
 PD 19-DEC-1996.
 XX
 PF 05-JUN-1996; 96WO-US09456.
 XX
 PR 07-JUN-1995; 95US-0480229.
 XX
 PA (PROG-) PROGENITOR INC.
 PA (UYVA-) UNIV VANDERBILT.
 XX
 PI Hogan B, Quertermous T, Snodgrass HR, Zupancic TJ;
 XX
 XX WPI; 1997-052233/05.
 DR N-PSDB; T47343.
 XX
 XX New developmentally regulated endothelial cell locus-1 (del-1) gene
 PT - used to develop prods. for the diagnosis and treatment of cancer
 PT and conditions involving abnormal angiogenesis
 XX
 PS Claim 4; Fig 6; 137pp; English.
 XX
 CC Human Del-1 (W10365) is the polypeptide product of the human
 CC developmentally-regulated endothelial cell locus-1 (del-1) gene
 CC (T47343). It shows 94% amino acid homology to the mouse Del-1
 CC protein (W10364). Structurally, members of this novel gene family
 CC contain 3 EGF-like domains and 2 discoidein I/factor VIII-like
 CC domains. Del-1 is expressed in endothelial and certain tumour
 CC cells. Its ability to inhibit vascular formation allows its used
 CC as an anti-angiogenic agent. It can be used as a tumour marker,
 CC to identify Del-1 binding partners, and to modulate endothelial
 CC cell growth and blood vessel formation. Recombinant Del-1 can be
 CC produced in transformed host cells utilising vectors incorporating
 CC del-1 nucleic acids.
 XX
 XX Sequence 481 AA;

Query Match 11.2%; Score 54; DB 18; Length 481;
 Best Local Similarity 100.0%; Pred. No. 2.6e-42;
 Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 341 FRTLNDMTWEPKARLDKGGKVNWTSCHNDQSQWLOVDLLVPTKVTGIIQ 394
 Db 342 ftrlndmftwepkarldkggkvnawtsghndsqwlvqvdllvptkvtgiltq 395

RESULT 9
 W10370
 ID W10370 standard; Protein; 102 AA.
 XX
 AC W10370;

DT 03-MAY-1997 (first entry)

DE Human Del-1 splice variant partial sequence.

XX Del-1; developmentally-regulated endothelial cell locus-1;
 KW signal transduction; cancer; tumour marker; angiogenesis;
 KW diabetic retinopathy; rheumatoid arthritis; endometriosis;
 KW cardiac ischaemia; stroke; vascular disease; wound healing;
 KW vulnery; bone formation; diagnosis; therapy.

XX Homo sapiens.

XX Key Location/Qualifiers
 FH 3..36
 FT Domain /label= EGF-1

FT Domain /note= "epidermal growth factor-like domain 1"
 FT 45..83
 FT /label= EGF-2
 FT /note= "epidermal growth factor-like domain 2"
 FT 90..102
 FT /label= EGF-3
 FT /note= "N-terminal portion of epidermal growth
 FT factor-like domain 3"
 XX
 XX W09640769-A1.
 PN
 XX 19-DEC-1996.
 PD
 XX 05-JUN-1996; 96WO-US09456.
 PF
 XX 07-JUN-1995; 95US-0480229.
 PR
 XX (PROG-) PROGENITOR INC.
 PA (UYVA-) UNIV VANDERBILT.
 XX
 XX Hogan B, Quertermous T, Snodgrass HR, Zupancic TJ;
 XX
 XX WPI; 1997-052233/05.
 DR
 XX New developmentally regulated endothelial cell locus-1 (del-1) gene
 PT - used to develop prods. for the diagnosis and treatment of cancer
 PT and conditions involving abnormal angiogenesis
 XX
 PS Example; Page 91; 137pp; English.
 XX
 CC A polypeptide sequence (W10370) shows a portion of a variant form
 CC of human developmentally-regulated endothelial cell locus-1 (del-1)
 CC polypeptide. In comparison with the major form (W10365) of human
 CC Del-1, 10 amino acid residues are missing between EGF-like domains
 CC 1 and 2 of the Del-1 variant. This is a result of alternative
 CC splicing (see also T47340).
 XX
 XX Sequence 102 AA;

Query Match 9.4%; Score 45; DB 18; Length 102;
 Best Local Similarity 100.0%; Pred. No. 1.8e-34;
 Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 PNPCHNGTCEISEAYRGDTFTGYVCKCPRGFNGHICQHINECE 124
 Db 47 pnpchngtceiseayrgdtftgyvckcprgfnghicqhinece 91

RESULT 10
 W94687
 ID W94687 standard; peptide; 43 AA.
 XX
 AC W94687;

DT 05-MAY-1999 (first entry)

DE Del-1 epidermal growth factor like domain #2.

XX Del-1; developmentally-regulated endothelial cell locus 1; cancer;
 KW discoidein I; factor VIII-like domain; epidermal growth factor; EGF;
 KW diabetic retinopathy; rheumatoid arthritis; endometriosis;
 KW angiogenesis.

XX Mus sp.

OS Synthetic.

XX US5877281-A.

PN

XX 02-MAR-1999.

PD

XX 05-JUN-1996; 96US-0659235.

PF

XX

```

PR 05-JUN-1996; 96US-0659235.
PR 07-JUN-1995; 95US-0480229.
XX (PROG-) PROGENITOR INC.
PA (UYVA-) UNIV VANDERBILT.
XX
XX Hogan B, Quertermous T, Snodgrass HR, Zupancic TJ;
XX WPI; 1999-189720/16.
XX
XX Del-1 polypeptide sequences - useful for treatment of cancer,
XX diabetic retinopathy, rheumatoid arthritis and endometriosis
XX
XX Claim 40; Fig 10; 73pp; English.
XX
XX The present sequence represents an epidermal growth factor like domain
XX from developmentally-regulated endothelial cell locus 1 (del-1). Del-1
XX also has discoidin I/factor VIII-like domains. The Del-1 proteins
XX have an inhibitory effect on angiogenesis (blood vessel growth), this
XX activity may be useful clinically to prevent neovascularisation of
XX tissues such as tumour nodules and prevention of metastases. The anti-
XX angiogenic activity of Del-1 may be used to treat abnormal conditions
XX that result from angiogenesis, including cancer, diabetic retinopathy,
XX rheumatoid arthritis and endometriosis. Since Del-1 promotes
XX angiogenesis it can be used to treat cardiac ischaemia, thrombotic
XX stroke, wound healing and peripheral vascular disease. Del-1 is also
XX useful for promoting bone formation. Del-1 binds to alpha V beta 3
XX integrin, and is an apoptosis inducer.
XX
XX Sequence 43 AA;
SQ
Query Match 8.5%; Score 41; DB 20; Length 43;
Best Local Similarity 100.0%; Pred. No. 4.6e-31;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 80 PNPCHNGTCETSEAYRGDTFTGYVCKCPRGFNIGHQHN1 120
Db 3 pnpchngtcetseayrgdtftgyvckcprgfnighcqh1 43
RESULT 11
W94697
ID W94697 standard; Protein; 321 AA.
XX
XX W94697;
XX
XX 05-MAY-1999 (first entry)
XX
XX Human milk fat globule protein MFG-E8.
XX
XX Del-1; developmentally-regulated endothelial cell locus 1; cancer;
XX discoidin I; factor VIII-like domain; epidermal growth factor; BGF;
XX diabetic retinopathy; rheumatoid arthritis; endometriosis;
XX angiogenesis.
XX
XX Homo sapiens.
XX
XX US5877281-A.
XX
XX 02-MAR-1999.
XX
XX 05-JUN-1996; 96US-0659235.
XX
XX 05-JUN-1996; 96US-0659235.
XX
XX 07-JUN-1995; 95US-0480229.
XX
XX (PROG-) PROGENITOR INC.
XX (UYVA-) UNIV VANDERBILT.
XX
XX Hogan B, Quertermous T, Snodgrass HR, Zupancic TJ;
XX WPI; 1999-189720/16.
XX
XX Del-1 polypeptide sequences - useful for treatment of cancer,
XX diabetic retinopathy, rheumatoid arthritis and endometriosis
XX
XX Example; Column 63-64; 73pp; English.
XX
XX The present sequence represents human milk fat globule protein MFG-E8,
XX which has homology to the developmentally-regulated endothelial cell
XX locus 1 (Del-1). The Del-1 protein has epidermal growth factor (EGF) like
XX domains and discoidin I/factor VIII-like domains. The Del-1 proteins
XX have an inhibitory effect on angiogenesis (blood vessel growth), this
XX activity may be useful clinically to prevent neovascularisation of
XX tissues such as tumour nodules and prevention of metastases. The anti-
XX angiogenic activity of Del-1 may be used to treat abnormal conditions
XX that result from angiogenesis, including cancer, diabetic retinopathy,
XX rheumatoid arthritis and endometriosis. Since Del-1 promotes
XX angiogenesis it can be used to treat cardiac ischaemia, thrombotic
XX stroke, wound healing and peripheral vascular disease. Del-1 is also
XX useful for promoting bone formation. Del-1 binds to alpha V beta 3
XX integrin, and is an apoptosis inducer.
XX
XX Sequence 321 AA;
SQ
Query Match 8.3%; Score 40; DB 20; Length 321;
Best Local Similarity 100.0%; Pred. No. 2.4e-29;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 341 FRTLNDMFTWEPRKARLDKQKVNAWTSGHNDQSOWLQV 380
Db 185 frtlndmftweprkarldkqgkvnawtsgndqsgwlqv 224
RESULT 12
W10368
ID W10368 standard; Peptide; 25 AA.
XX
XX W10368;
XX
XX 03-MAY-1997 (first entry)
XX
XX Human Del-1 signal peptide.
XX
XX Del-1; developmentally-regulated endothelial cell locus-1;
XX signal transduction; cancer; tumour marker; angiogenesis;
XX diabetic retinopathy; rheumatoid arthritis; endometriosis;
XX cardiac ischaemia; stroke; vascular disease; wound healing;
XX vulnerary; bone formation; diagnosis; therapy.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Peptide 3..21
XX /label= sig_peptide
XX
XX W09640769-A1.
XX
XX 19-DEC-1996.
XX
XX 05-JUN-1996; 96WO-US09456.
XX
XX 07-JUN-1995; 95US-0480229.
XX
XX (PROG-) PROGENITOR INC.
XX (UYVA-) UNIV VANDERBILT.
XX
XX Hogan B, Quertermous T, Snodgrass HR, Zupancic TJ;
XX WPI; 1997-052233/05.
XX
XX New developmentally regulated endothelial cell locus-1 (del-1) gene
XX - used to develop prods. for the diagnosis and treatment of cancer
XX and conditions involving abnormal angiogenesis

```

```

XX PS Example: Fig 9; 137pp; English.
XX CC The predicted amino acid sequence (W10368) at the N-terminus
XX CC of the human developmentally-regulated endothelial cell locus-1
XX CC (del-1) gene product (W10365) shows characteristics common to
XX CC signal peptides. The putative signal begins with a basic Arg
XX CC residue and is followed by a stretch of 18 hydrophobic residues.
XX CC The Chou and Fasman algorithm predicts that the putative signal
XX CC sequence is followed by a protein turn structure, a feature
XX CC commonly found after signal peptides. The Del-1 protein is
XX CC secreted by expressing cells.
XX SQ Sequence 25 AA;

Query Match 3.8%; Score 18; DB 18; Length 25;
Best Local Similarity 100.0%; Pred. No. 9.7e-10;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 WLLVGLSLGVPOFGKGDI 25
Db 8 WLLVGLSLGVPOFGKGDI 25

RESULT 13
W94698
ID W94698 standard; peptide; 25 AA.
AC W94698;
XX DT 05-MAY-1999 (first entry)
XX DE Human Del-1 signal protein.
XX KW Del-1; developmentally-regulated endothelial cell locus 1; cancer;
XX KW discoidin I; factor VIII-like domain; epidermal growth factor; EGF;
XX KW diabetic retinopathy; rheumatoid arthritis; endometriosis;
XX OS Homo sapiens.
XX PN US5877281-A.
XX PD 02-MAR-1999.
XX PF 05-JUN-1996; 96US-0659235.
XX PR 05-JUN-1996; 96US-0659235.
XX PR 07-JUN-1995; 95US-0480229.
XX PA (PROG-) PROGENITOR INC.
XX PA (UYVA-) UNIV VANDERBILT.
XX PI Hogan B, Quertermous T, Snodgrass HR, Zupancic TJ;
XX WPI; 1999-189720/16.
XX DR Del-1 polypeptide sequences - useful for treatment of cancer,
XX PT diabetic retinopathy, rheumatoid arthritis and endometriosis
XX PS Example: Column 63-65; 73pp; English.
XX CC The present sequence is the human developmentally-regulated endothelial
XX CC cell locus 1 (Del-1) amino terminus peptide sequence. The Del-1 protein
XX CC has epidermal growth factor (EGF) like domains and discoidin I/factor
XX CC VIII-like domains. The Del-1 proteins have an inhibitory effect on
XX CC angiogenesis (blood vessel growth), this activity may be useful
XX CC clinically to prevent neovascularisation of tissues such as tumour
XX CC nodules and prevention of metastases. The anti-angiogenic activity of
XX CC Del-1 may be used to treat abnormal conditions that result from
XX CC angiogenesis, including cancer, diabetic retinopathy, rheumatoid
XX CC arthritis and endometriosis. Since Del-1 promotes angiogenesis it can be
us-09-237-981-10.oligo.rag
CC used to treat cardiac ischaemia, thrombotic stroke, wound healing and
CC peripheral vascular disease. Del-1 is also useful for promoting bone
CC formation. Del-1 binds to alpha V beta 3 integrin, and is an apoptosis
CC inducer.
XX SQ Sequence 25 AA;

Query Match 3.8%; Score 18; DB 20; Length 25;
Best Local Similarity 100.0%; Pred. No. 9.7e-10;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 WLLVGLSLGVPOFGKGDI 25
Db 8 WLLVGLSLGVPOFGKGDI 25

RESULT 14
W94688
ID W94688 standard; peptide; 42 AA.
XX AC W94688;
XX DT 05-MAY-1999 (first entry)
XX DE Del-1 epidermal growth factor like domain #3.
XX KW Del-1; developmentally-regulated endothelial cell locus 1; cancer;
XX KW discoidin I; factor VIII-like domain; epidermal growth factor; EGF;
XX KW diabetic retinopathy; rheumatoid arthritis; endometriosis;
XX KW angiogenesis.
XX OS Mus sp.
XX OS Synthetic.
XX PN US5877281-A.
XX PD 02-MAR-1999.
XX PF 05-JUN-1996; 96US-0659235.
XX PR 05-JUN-1996; 96US-0659235.
XX PR 07-JUN-1995; 95US-0480229.
XX PA (PROG-) PROGENITOR INC.
XX PA (UYVA-) UNIV VANDERBILT.
XX PI Hogan B, Quertermous T, Snodgrass HR, Zupancic TJ;
XX WPI; 1999-189720/16.
XX DR Del-1 polypeptide sequences - useful for treatment of cancer,
XX PT diabetic retinopathy, rheumatoid arthritis and endometriosis
XX PS Claim 40; Fig 10; 73pp; English.
XX CC The present sequence represents an epidermal growth factor like domain
XX CC from developmentally-regulated endothelial cell locus 1 (Del-1). Del-1
XX CC also has discoidin I/factor VIII-like domains. The Del-1 proteins
XX CC have an inhibitory effect on angiogenesis (blood vessel growth), this
XX CC activity may be useful clinically to prevent neovascularisation of
XX CC tissues such as tumour nodules and prevention of metastases. The anti-
XX CC angiogenic activity of Del-1 may be used to treat abnormal conditions
XX CC that result from angiogenesis, including cancer, diabetic retinopathy,
XX CC rheumatoid arthritis and endometriosis. Since Del-1 promotes
XX CC angiogenesis it can be used to treat cardiac ischaemia, thrombotic
XX CC stroke, wound healing and peripheral vascular disease. Del-1 is also
XX CC useful for promoting bone formation. Del-1 binds to alpha V beta 3
XX CC integrin, and is an apoptosis inducer.
XX SQ Sequence 42 AA;

```

Query Match 3.8%; Score 18; DB 20; Length 42;
 Best Local Similarity 100.0%; Pred. No. 1.5e-09;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 140 NYSCECPGEFGMGRNCQYK 157
 DB 25 nyscecpgefgmgnrcqyk 42

RESULT 15
 W94686
 ID W94686 standard; peptide; 57 AA.
 AC W94686;
 XX
 DT 05-MAY-1999 (first entry)
 XX
 DE Del-1 epidermal growth factor like domain #1.
 XX
 KW Del-1; developmentally-regulated endothelial cell locus 1; cancer;
 KW discoidin 1; factor VIII-like domain; epidermal growth factor; EGF;
 KW diabetic retinopathy; rheumatoid arthritis; endometriosis;
 KW angiogenesis.
 XX
 OS Mus sp.
 OS Synthetic.
 XX
 PN US5877281-A.
 XX
 PD 02-MAR-1999.
 XX
 PF 05-JUN-1996; 96US-0659235.
 XX
 PR 05-JUN-1996; 96US-0659235.
 PR 07-JUN-1995; 95US-0480229.
 XX
 PA (PROG-) PROGENITOR INC.
 PA (UYVA-) UNIV VANDERBILT.
 XX
 PI Hogan B, Quettermous T, Snodgrass HR, Zupancic TJ;
 XX WPI; 1999-189720/16.
 DR
 XX
 PT Del-1 polypeptide sequences - useful for treatment of cancer,
 PT diabetic retinopathy, rheumatoid arthritis and endometriosis
 XX
 PS Claim 40; Fig 10; 73pp; English.
 XX
 CC The present sequence represents an epidermal growth factor like domain
 CC from developmentally-regulated endothelial cell locus 1 (Del-1). Del-1
 CC also has discoidin I/factor VIII-like domains. The Del-1 proteins
 CC have an inhibitory effect on angiogenesis (blood vessel growth), this
 CC activity may be useful clinically to prevent neovascularization of
 CC tissues such as tumour nodules and prevention of metastases. The anti-
 CC angiogenic activity of Del-1 may be used to treat abnormal conditions
 CC that result from angiogenesis, including cancer, diabetic retinopathy,
 CC rheumatoid arthritis and endometriosis. Since Del-1 promotes
 CC angiogenesis it can be used to treat cardiac ischaemia, thrombotic
 CC stroke, wound healing and peripheral vascular disease. Del-1 is also
 CC useful for promoting bone formation. Del-1 binds to alpha V beta 3
 CC integrin, and is an apoptosis inducer.
 XX
 SQ Sequence 57 AA;

Query Match 2.9%; Score 14; DB 20; Length 57;
 Best Local Similarity 100.0%; Pred. No. 1.1e-05;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 PNCSSVVEVASDDE 70
 DB 37 pncssvvevasdee 50

RESULT 16
 W87896
 ID W87896 standard; Peptide; 1010 AA.
 XX
 AC W87896;
 XX
 DT 26-APR-1999 (first entry)
 XX
 DE Human JAGGED1 soluble polypeptide.
 XX
 KW JAGGED; JAGGED1; hJAGGED1; human; notch ligand; stem cell;
 KW progenitor cell; haematopoiesis; cell differentiation;
 KW Alagille syndrome; leukaemia; lymphoma; diagnosis; therapy.
 XX
 OS Homo sapiens.
 XX
 FH Key
 FT Peptide
 FT /note= "signal peptide"
 FT Protein
 FT /note= "mature protein"
 FT Domain
 FT /note= "Delta/Serrate/Lag-2 (DSL) domain"
 FT Region
 FT /note= "EGF-like repeat region"
 XX
 PN W09858958-A2.
 XX
 PD 30-DEC-1998.
 XX
 PF 25-JUN-1998; 98WO-US13207.
 XX
 PR 25-JUN-1997; 97US-0882046.
 XX
 PA (CHIL-) CHILDRENS HOSPITAL PHILADELPHIA.
 PA (UNIW) UNIV WASHINGTON.
 XX
 PI Hood L, Krantz ID, Li L, Spinner NB;
 XX WPI; 1999-081220/07.
 DR
 XX
 PT New Jagged peptides for inhibiting differentiation of progenitor
 PT cells - also used for maintaining these cells in undifferentiated
 PT state, e.g. for haematopoietic reconstitution
 XX
 PS Claim 8; Page -; 101pp; English.
 XX
 CC This is the amino acid sequence of a biologically active soluble
 CC human JAGGED1 (hJAGGED) polypeptide comprising amino acid residues
 CC 1-1010 of hJAGGED1 (see W87894). It was prepared by PCR
 CC amplification (see V63759-60) of hJAGGED1 cDNA (see V63753) and
 CC expression in CHO and BHK cells. hJAGGED1 is an activating ligand
 CC for Notch protein that is expressed in bone marrow stromal cells.
 CC A stromal cell line expressing hJAGGED1 permits survival and
 CC proliferation of haematopoietic progenitor cells expressing Notch
 CC but inhibits granulocyte differentiation. hJAGGED1 and active
 CC peptides can be used (i) to inhibit differentiation of haematopoietic
 CC progenitor cells (HPC), e.g. for subsequent production of blood cells
 CC for transplantation or dendritic cells for immunotherapy, and (ii) to
 CC maintain HPC in the undifferentiated state, particularly totipotent
 CC cells or cells able to reconstitute the haematopoietic system, e.g.
 CC in patients with leukaemia or lymphoma. Treated HPC, e.g. where
 CC taken from a neonate, may be cryopreserved for many years, then
 CC thawed for further expansion and differentiation. Optionally
 CC JAGGED polypeptides are provided by the transformed host cells.
 CC (N.B. the amino acid sequence of the polypeptide of W87896 was
 CC constructed from the full-length hJAGGED1 amino acid sequence given
 CC in Fig1B of the specification).
 XX
 SQ Sequence 1010 AA;

Query Match 2.1%; Score 10; DB 20; Length 1010;
Best Local Similarity 100.0%; Pred. No. 0.79; 0; Indels 0; Gaps 0;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 80 PNPCHNGGTC 89
Db 750 pnpchnggctc 759
RESULT 17
W18351
ID W18351 standard; protein; 1036 AA.
XX
AC W18351;
XX
DT 11-FEB-1998 (first entry)
XX
DE Proliferation and differentiation suppression polypeptide.
XX
KW Proliferation; differentiation; suppression; human; delta-1;
KW serrate-1; blood cell; neuron; leukaemia; malignant tumour;
KW immunosuppression.
OS Homo sapiens.
XX
PN W09719172-A1.
XX
PD 29-MAY-1997.
XX
PF 15-NOV-1996; 96WO-JP033356.
XX
PR 30-NOV-1995; 95JP-0311811.
XX
PR 17-NOV-1995; 95JP-0299611.
XX
XX (ASAH) ASAH KASEI KOGYO KK.
XX
XX Itoh A, Sakano S;
XX
XX WPI; 1997-298110/27.
XX
XX Peptide(s) encoded by human genes delta-1 and serrate-1 - suppress
PT proliferation and differentiation of undifferentiated human blood
PT cells
XX
XX Claim 5; Page 66-71; 114pp; Japanese.
XX
XX The present sequence represents a polypeptide which suppresses
CC proliferation and differentiation of undifferentiated cells such
CC as neurons and blood cells. The polypeptide may be used for the
CC prevention and control of disorders involving undifferentiated
CC cells, such as leukaemia and malignant tumours, and improvement of
CC blood formation, e.g. after immunosuppression.
XX
XX Sequence 1036 AA;
SQ
Query Match 2.1%; Score 10; DB 18; Length 1036;
Best Local Similarity 100.0%; Pred. No. 0.81;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 80 PNPCHNGGTC 89
Db 719 pnpchnggctc 728
RESULT 18
W18352
ID W18352 standard; protein; 1187 AA.
XX
AC W18352;
XX
DT 11-FEB-1998 (first entry)
XX
XX

DE Proliferation and differentiation suppression polypeptide.
XX
KW Proliferation; differentiation; suppression; human; delta-1;
KW serrate-1; blood cell; neuron; leukaemia; malignant tumour;
KW immunosuppression.
OS Homo sapiens.
XX
PN W09719172-A1.
XX
PD 29-MAY-1997.
XX
PF 15-NOV-1996; 96WO-JP033356.
XX
PR 30-NOV-1995; 95JP-0311811.
XX
PR 17-NOV-1995; 95JP-0299611.
XX
XX (ASAH) ASAH KASEI KOGYO KK.
XX
XX Itoh A, Sakano S;
XX
XX WPI; 1997-298110/27.
XX
XX Peptide(s) encoded by human genes delta-1 and serrate-1 - suppress
PT proliferation and differentiation of undifferentiated human blood
PT cells
XX
XX Claim 6; Page 71-76; 114pp; Japanese.
XX
XX The present sequence represents a polypeptide which suppresses
CC proliferation and differentiation of undifferentiated cells such
CC as neurons and blood cells. The polypeptide may be used for the
CC prevention and control of disorders involving undifferentiated
CC cells, such as leukaemia and malignant tumours, and improvement of
CC blood formation, e.g. after immunosuppression.
XX
XX Sequence 1187 AA;
SQ
Query Match 2.1%; Score 10; DB 18; Length 1187;
Best Local Similarity 100.0%; Pred. No. 0.92; 0; Indels 0; Gaps 0;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 80 PNPCHNGGTC 89
Db 719 pnpchnggctc 728
RESULT 19
W05835
ID W05835 standard; Protein; 1193 AA.
XX
AC W05835;
XX
DT 28-JAN-1997 (first entry)
XX
DE Chick Serrate.
XX
XX C-Serrate; Notch; cell differentiation; cell fate; tissue repair;
KW central nervous system; cancer; therapy; diagnosis.
XX
XX Gallus sp.
XX
XX Key Location/Qualifiers
FH Domain 1..1041
FT /label= Extracellular_domain
FT Peptide 1..5
FT /label= Sig_peptide
FT /note= "lacks the N-terminal portion owing to
FT truncation of the encoding cDNA clone"
FT Domain 158..203
FT /label= DSL
FT /note= "region of homology with Drosophila Delta

FT and Serrate, predicted to mediate binding
FT with Notch"
FT 208..837
FT /label= ELR
FT /note= "epidermal growth factor-like repeat domain"
FT 208..238
FT /label= ELR1
FT 239..274
FT /label= ELR2
FT 275..313
FT /label= ELR3
FT 314..351
FT /label= ELR4
FT 352..390
FT /label= ELR5
FT 391..427
FT /label= ELR6
FT 428..464
FT /label= ELR7
FT 465..502
FT /label= ELR8
FT 503..540
FT /label= ELR9
FT 541..606
FT /label= ELR10
FT 607..644
FT /label= ELR11
FT 655..682
FT /label= ELR12
FT 683..721
FT /label= ELR13
FT 722..759
FT /label= ELR14
FT 760..797
FT /label= ELR15
FT 798..837
FT /label= ELR16
FT 854..911
FT /label= Cysteine-rich_region
FT 1042..1066
FT /label= Transmembrane_domain
FT 1067..1193
FT /label= Intracellular_domain
XX
XX
PN W09627610-Al.
XX
PD 12-SEP-1996.
XX
XX 07-MAR-1996; 96WO-US03172.
XX
XX 07-MAR-1995; 95US-0400159.
XX (IMCR) IMPERIAL CANCER RES TECHNOLOGY.
XX (UYA) UNIV YALE.
XX
XX Artavanis-Tsakonas S, Gray GE, Henrique DMP, Ish-Horowitz D;
XX Lewis JH, Mann RS, Myat AM;
XX WPI; 1996-425379/42.
XX N-PSDB; T40092.
XX
XX Vertebrate Serrate protein and related DNA - used to treat or
XX prevent malignancies characterised by increased Notch activity.
XX
XX Disclosure; Page 112-115; 161pp; English.
XX
XX Chicken Serrate (W05835), or C-Serrate, is a ligand for the zygotic
XX neurogenic locus Notch and is believed to play a major role in
XX determining cell fates in the central nervous system. Its amino
XX acid sequence was deduced from a cDNA clone (T40092) obtd. from an
XX optic explant cDNA library. C-Serrate is expressed in the central
XX nervous system, cranial placodes, nephric mesoderm, vascular
XX system, and limb bud mesenchyme.

XX
SQ Sequence 1193 AA;

Query Match 2.1%; Score 10; DB 17; Length 1193;
Best Local Similarity 100.0%; Pred. No. 0.92;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 PNPCHNGGTC 89
|||||
Db 724 pnpchngtc 733

RESULT 20
Y59599
ID Y59599 standard; Protein; 1193 AA.
XX AC Y59599;
XX
XX 05-APR-2000 (first entry)
XX
DE Chick Serrate protein sequence.
XX
KW Serrate; cancerous condition; melanoma; lung cancer; breast cancer;
KW nervous system disorder; infection; nutritional disease; therapy;
KW cell proliferation promoter; tissue regeneration; chicken.
XX
OS Gallus sp.
XX
XX US6004924-A.
XX
XX 21-DEC-1999.
XX
XX 06-MAR-1996; 96US-0611729.
XX
PR 11-DEC-1991; 91US-0808458.
PR 14-SEP-1993; 93US-0121979.
PR 07-JUN-1994; 94US-0255102.
PR 07-MAR-1995; 95US-0400159.
XX
XX (IMCR) IMPERIAL CANCER RES TECHNOLOGY.
XX (UYA) UNIV YALE.
XX
XX Ish-Horowitz D, Henrique DMP, Myat AM, Fleming RJ;
XX Artavanis-Tsakonas S, Gray GE, Mann RS, Lewis JH;
XX WPI; 2000-105089/09.
XX N-PSDB; Z49098.
XX
XX Purified Serrate proteins useful for treating neoplasias, nervous
XX disorders and for promoting cell proliferation and tissue regeneration
XX and repair -
XX
XX Claim 1; Fig 12; 114pp; English.
XX
XX This sequence represents a chick serrate protein.
XX The invention relates to purified vertebrate (mouse, chick, and human)
XX Serrate proteins. The Serrate proteins may be administered to treat a
XX cancerous condition (e.g. melanoma, lung cancer and breast cancer) by
XX preventing progression from a preneoplastic or nonmalignant state into a
XX neoplastic or malignant state. It may also be used to treat nervous
XX system disorders (such as lesions caused by infections, nutritional
XX disease and toxic substances) and to promote cell proliferation and
XX tissue regeneration and repair. The protein itself is administered to
XX supplement a patient's own production of Serrate proteins (if levels of
XX expression are low) or to compensate for expression of inactive proteins
XX due to genetic mutations. The protein may also be used in the production
XX of antibodies against Serrate proteins which may be used to either down
XX regulate Serrate activity or to detect Serrate proteins in samples (for
XX example via enzyme-linked immunosorbant assay (ELISA)). The proteins may
XX also be used to study Serrate expression and its role in metabolism and
XX to assay for agents which modulate its expression and activity.
XX

SQ Sequence 1193 AA;

Query Match 2.1%; Score 10; DB 21; Length 1193;
Best Local Similarity 100.0%; Pred. No. 0.92;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 PNPCHNGGTC 89
Db 724 pnpchnggtc 733

RESULT 21

W40827 ID W40827 standard; Protein; 1208 AA.

XX W40827;

DT 21-MAY-1998 (first entry)

XX Human Jagged protein.

KW Jagged; Notch; angiogenesis; endothelial cell; migration; human;
KW wound repair; vulnery; injury repair; signal transduction;
KW motor neurone disease; amyotrophic lateral sclerosis; polymyelitis;
KW diagnosis; therapy.

XX Homo sapiens.

PH Key Location/Qualifiers

FT Peptide I..11

FT /label= Sig_peptide

FT Domain 175..220

FT /note= "DSL (Delta, Serrate, Lag-2 and Apx-1)
FT domain"

FT Region 224..852

FT /note= "EGF-like repeat region containing 16
FT EGF repeats"

FT Misc-difference 526

FT /note= "encoded by ANC"

FT Region 853..992

FT /note= "cysteine-rich region"

FT Domain 1058..1083

FT /note= "transmembrane domain"

FT Region 1084..1208

FT /note= "cytoplasmic region"

XX W09745143-A1.

XX 04-DEC-1997.

XX 30-MAY-1997; 97WO-US09407.

XX 31-MAY-1996; 96US-0018841.

XX (NAAM-) NAT AMERICAN RED CROSS.

XX (UYGE-) UNIV GENEVE.

XX Maciag T, Montesano R, Pepper M, Wong MK, Zimrin AB;

XX WPI; 1998-032340/03.

XX N-PSDB; W03674.

XX New human Jagged protein - used to inhibit or promote angiogenesis
XX and to control migration of endothelial cells in injured blood
XX vessels
XX Claim 2; Page 54-61; 81pp; English.

XX This sequence comprises the human homologue of the rat Jagged
XX 'protein. Jagged is able to bind Notch protein and is involved in
XX endothelial cell (EC) migration and differentiation. The human
XX Jagged amino acid sequence was deduced from a human endothelial

CC cell cDNA (see W03674) induced by exposure to fibrin. Jagged
CC polypeptides can be expressed in host cell systems. A method for
CC treating or preventing disease by administering an agent that
CC (ant)agonises, inhibits, prevents, enhances or stimulates function
CC of the Notch or Jagged proteins is claimed, as well as a method for
CC affecting differentiation of mesoderm, endoderm, ectoderm and/or
CC neuroderm cells. When Jagged is applied to a micro-diameter blood
CC vessel from which ECs have been removed, damaged or reduced, it
CC decrease migrations of EC to the site, but when delivered to a
CC similar site on a large vessel it increases EC migration. Jagged
CC and its agonists are used to inhibit or prevent angiogenesis (where
CC associated with solid tumours, rheumatoid arthritis, inflammation,
CC or restenosis, particularly preventing angiogenesis from the vaso
CC vasorum and promoting large vessel EC migration to repair the lumen
CC of large vessels). Anti-Jagged and Jagged antagonists (e.g.
CC antisense Jagged and Jagged mutants) are used to promote or enhance
CC angiogenesis, particularly for wound and injury repair, e.g. where
CC surgical, traumatic and/or caused by disease, e.g. diabetes-related
CC (all claimed). Angiogenesis can be modulated in vitro or in vivo
CC and expression of proteins by gene therapy is included. Modulation
CC of the Notch-Jagged signalling pathway may also be involved in
CC placental development and motor neurone diseases such as
CC amyotrophic lateral sclerosis, poliomyelitis etc.

XX Sequence 1208 AA;

Query Match 2.1%; Score 10; DB 19; Length 1208;

Best Local Similarity 100.0%; Pred. No. 0.93;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 PNPCHNGGTC 89

Db 740 pnpchnggtc 749

RESULT 22

W05833 ID W05833 standard; Protein; 1218 AA.

XX W05833;

XX 28-JAN-1997 (first entry)

XX Human Serrate-1 (HJ1).

XX Serrate-1; human jagged-1; HJ1; Notch; cell differentiation;

XX cell fate; central nervous system; cancer; tissue repair; therapy;

XX diagnosis; antibody.

XX Homo sapiens.

XX Key Location/Qualifiers

XX Domain 1..1067

XX /label= Extracellular_domain

XX Peptide 14..29

XX /label= Sig_peptide

XX Domain 185..229

XX /label= DSL

XX /note= "region of homology with Drosophila Delta

XX and Serrate, predicted to mediate binding

XX with Notch"

XX Domain 234..896

XX /label= ELR

XX /note= "epidermal growth factor-like repeat domain"

XX Region 234..264

XX /label= ELR1

XX Region 265..299

XX /label= ELR2

XX Region 300..339

XX /label= ELR3

XX Region 340..377

XX /label= ELR4

FT Region 378..415
FT /label= ELR5
FT Region 416..453
FT /label= ELR6
FT Region 454..490
FT /label= ELR7
FT Region 491..528
FT /label= ELR8
FT Region 529..566
FT /label= ELR9
FT Region 567..598
FT /label= Partial_ELR
FT Region 599..632
FT /label= Partial_ELR
FT Region 633..670
FT /label= ELR10
FT Region 671..708
FT /label= ELR11
FT Region 709..747
FT /label= ELR12
FT Region 748..785
FT /label= ELR13
FT Region 786..823
FT /label= ELR14
FT Region 824..862
FT /label= ELR15
FT Region 863..879
FT /label= Partial_ELR
FT Region 880..896
FT /label= Partial_ELR
FT Domain 1068..1089
FT /label= Transmembrane_domain
FT Domain 1090..1218
FT /label= Intracellular_domain
XX
PN W09627610-A1.
XX
XX
PD 12-SEP-1996.
XX
XX 07-MAR-1996; 96WO-US03172.
XX
XX 07-MAR-1995; 95US-0400159.
XX
XX (IMCR) IMPERIAL CANCER RES TECHNOLOGY.
XX PA (UYVA) UNIV YALE.
XX
XX Artavanis-Tsakonas S, Gray GE, Henrique DMP, Ish-Horowicz D;
XX PI Lewis JH, Mann RS, Myat AM;
XX
XX WPI; 1996-425379/42.
XX DR N-PSDB; T40090.
XX
XX Vertebrate Serrate protein and related DNA - used to treat or
XX PT prevent malignancies characterised by increased Notch activity.
XX
XX Claim 4; Page 95-98; 161pp; English.
XX
XX Human Serrate-1 (W05833) and human Serrate-2 (W05833) are ligands
XX CC for the zygotic neurogenic locus Notch, and are believed to play a
XX CC major role in determining cell fates (differentiation) in the
XX CC central nervous system. Their amino acid sequences were deduced
XX CC from cDNA clones (see also T40090-91) isolated from human foetal
XX CC brain cDNA libraries. The proteins, antibodies raised to them,
XX CC and encoding nucleic acids can be used in the detection of
XX CC Serrate sequences and in the treatment of disorders of cell fate
XX CC or differentiation, partic. cancer, nervous system disorders
XX CC and in tissue repair or regeneration.
XX
SQ Sequence 1218 AA;

Query Match 2.18; Score 10; DB 17; Length 1218;
Best Local Similarity 100.0%; Pred. No. 0.94;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 80 PNPCHNGGTC 89
Db 750 pnpchnggtc 759
|||||
RESULT 23
W18354
ID W18354 standard; protein; 1218 AA.
XX
XX W18354;
XX
XX 11-FEB-1998 (first entry)
XX Proliferation and differentiation suppression polypeptide.
XX
XX
XX Proliferation; differentiation; suppression; human; delta-1;
XX serrate-1; blood cell; neuron; leukaemia; malignant tumour;
XX immunosuppression.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX FT Peptide 1..31
XX FT /label= Signal
XX FT Protein 32..1218
XX FT /label= Differentiation_suppression_protein
XX
XX W09719172-A1.
XX
XX 29-MAY-1997.
XX
XX 15-NOV-1996; 96WO-JP03356.
XX
XX 30-NOV-1995; 95JP-0311811.
XX PR 17-NOV-1995; 95JP-0299611.
XX
XX (ASAH) ASahi KASEI KOGYO KK.
XX
XX Itoh A, Sakano S;
XX WPI; 1997-298110/27.
XX DR N-PSDB; T70175.
XX
XX Peptide(s) encoded by human genes delta-1 and serrate-1 - suppress
XX PT proliferation and differentiation of undifferentiated human blood
XX PT cells
XX
XX Claim 15; Page 83-91; 114pp; Japanese.
XX
XX The present sequence represents a polypeptide which suppresses
XX CC proliferation and differentiation of undifferentiated cells such
XX CC as neurons and blood cells. The polypeptide may be used for the
XX CC prevention and control of disorders involving undifferentiated
XX CC cells, such as leukaemia and malignant tumours, and improvement of
XX CC blood formation, e.g. after immunosuppression.
XX
XX Sequence 1218 AA;
Query Match 2.18; Score 10; DB 18; Length 1218;
Best Local Similarity 100.0%; Pred. No. 0.94;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 80 PNPCHNGGTC 89
Db 750 pnpchnggtc 759
|||||
RESULT 24
W44301
ID W44301 standard; protein; 1218 AA.

```

XX AC W44301;
XX DT 19-JUN-1998 (first entry)
XX DE Human serrate 1.
XX KW Human; serrate 2; regulation; stem cell; differentiation; neoplasm;
KW leukaemia; endothelial cell; tumour.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT Peptide 1..31
XX FT Protein /label= Signal
XX FT /label= Serrate-1
XX PN W09802458-A1.
XX XX
XX PD 22-JAN-1998.
XX PF 11-JUL-1997; 97WO-JP02414.
XX PR 14-MAY-1997; 97JP-0124063.
XX PR 16-JUL-1996; 96JP-0186220.
XX XX
XX PA (ASAH ) ASahi KASEI KOGYO KK.
XX PI Itoh A, Sakano S;
XX XX
XX DR WPI; 1998-110528/10.
XX DR N-PSDB; V15201.
XX XX
XX PT Human serrate-2 gene expression products - used to regulate stem
XX PT cell differentiation, useful in treating neoplasms, e.g. leukaemia
XX PS Disclosure; Page 77-86; 103pp; Japanese.
XX CC The present sequence represents human serrate 1, from the present
XX CC invention which describes human serrate 2. The present invention also
XX CC describes a method for the preparation of the polypeptides, and
XX CC antibodies binding to the polypeptide and its fragments. The polypeptide
XX CC and its fragments expressed by the serrate-2-gene can be used to inhibit
XX CC stem (especially blood stem) cell differentiation and to inhibit
XX CC endothelial cell growth. They may be incorporated in a cell culture
XX CC media for culturing undifferentiated stem cells. They can also be used
XX CC for treatment of neoplasms such as leukaemia. The antibodies can be used
XX CC for the diagnosis of malignant tumours.
XX SQ Sequence 1218 AA;

Query Match 2.1%; Score 10; DB 19; Length 1218;
Best Local Similarity 100.0%; Pred. No. 0.94;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 PNPCHNGGTC 89
Db 750 pnpchnggtc 759
|||||

RESULT 25
W87894
ID W87894 standard; Protein; 1218 AA.
XX AC W87894;
XX DT 26-APR-1999 (first entry)
XX DE Human JAGGED1 protein.
XX KW JAGGED; JAGGED1; hJAGGED1; human; notch ligand; stem cell;

```

```

KW KW progenitor cell; haematopoiesis; cell differentiation;
KW Alagille syndrome; leukaemia; lymphoma; diagnosis; therapy.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT Peptide 1..21
XX FT Protein /note= "signal peptide"
XX FT /note= "mature protein"
XX FT Domain 185..239
XX FT /note= "Delta/Serrate/Lag-2 (DSL) domain"
XX FT Region 234..862
XX FT /note= "EGF-like repeat region"
XX FT Region 863..1012
XX FT /note= "cysteine-rich region"
XX FT Domain 1077..1091
XX FT /note= "transmembrane domain"
XX FT Peptide 188..204
XX FT /note= "this peptide is specifically claimed in
XX FT Claim 1"
XX FT Peptide 178..240
XX FT /note= "this soluble peptide is specifically
XX FT claimed in Claim 8"
XX FT Protein 1..1010
XX FT /note= "this soluble protein is specifically
XX FT claimed in Claim 8"
XX XX
XX PN W09858958-A2.
XX XX
XX PD 30-DEC-1998.
XX XX
XX PF 25-JUN-1998; 98WO-US13207.
XX XX
XX PR 25-JUN-1997; 97US-0882046.
XX XX
XX PA (CHIL-) CHILDRENS HOSPITAL PHILADELPHIA.
XX PA (UNIW ) UNIV WASHINGTON.
XX XX
XX PI Hood L, Krantz ID, Li L, Spinner NB;
XX XX
XX DR WPI; 1999-081220/07.
XX DR N-PSDB; V63753.
XX XX
XX PT New Jagged peptides for inhibiting differentiation of progenitor
XX PT cells - also used for maintaining these cells in undifferentiated
XX PT state, e.g. for haematopoietic reconstitution
XX PS Claim 6; Fig 1A; 101pp; English.
XX XX
XX CC This is the amino acid sequence of human JAGGED1 (hJAGGED1), an
XX CC activating ligand for Notch protein. hJAGGED1 is expressed in
XX CC bone marrow stromal cells, and a stromal cell line expressing
XX CC hJAGGED1 permits survival and proliferation of haematopoietic
XX CC progenitor cells expressing Notch but inhibits granulocyte
XX CC differentiation. A cDNA clone (see V63753) encoding hJAGGED1
XX CC was obtained from a human bone marrow cDNA library. hJAGGED1
XX CC polypeptides and biologically active peptides (see W87896-98) are
XX CC able (i) to inhibit differentiation of haematopoietic progenitor
XX CC cells (HPC), e.g. for subsequent production of blood cells for
XX CC transplantation or dendritic cells for immunotherapy, and (ii) to
XX CC maintain HPC in the undifferentiated state, particularly totipotent
XX CC cells or cells able to reconstitute the haematopoietic system, e.g.
XX CC in patients with leukaemia or lymphoma. Treated HPC, e.g. where
XX CC taken from a neonate, may be cryopreserved for many years, then
XX CC thawed for further expansion and differentiation. Optionally
XX CC Jagged may be provided by cells transformed to express the
XX CC membrane-bound protein. Antibodies raised against hJAGGED1 can
XX CC be used in a method of diagnosing Alagille syndrome by detecting
XX CC reduced expression of hJAGGED1 or expression of abnormal hJAGGED1.
XX SQ Sequence 1218 AA;

```

Query Match 2.1%; Score 10; DB 20; Length 1218;
Best Local Similarity 100.0%; Pred. No. 0.94;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 PNPCHNGGTC 89
 |||||
Db 750 popchnggtc 759

Search completed: May 23, 2001, 06:21:18
Job time: 437 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 06:19:06 ; Search time 76.77 Seconds
(without alignments)
1005.734 Million cell updates/sec

Title: US-09-237-981-10
Perfect score: 480
Sequence: 1 MKHLVAALLVGLSLGVPQF.....WSWYGRITLRSLLGCAEEE 480

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 1009251 seqs, 160854530 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1009251

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 50 summaries

Database : Pending_Patents_AA_Main:*
1: /cgn2_6/ptodata/2/paa/PCTUS_COMB.pep.*
2: /cgn2_6/ptodata/2/paa/US06_COMB.pep.*
3: /cgn2_6/ptodata/2/paa/US07_COMB.pep.*
4: /cgn2_6/ptodata/2/paa/US080_COMB.pep.*
5: /cgn2_6/ptodata/2/paa/US081_COMB.pep.*
6: /cgn2_6/ptodata/2/paa/US082_COMB.pep.*
7: /cgn2_6/ptodata/2/paa/US083_COMB.pep.*
8: /cgn2_6/ptodata/2/paa/US084_COMB.pep.*
9: /cgn2_6/ptodata/2/paa/US085_COMB.pep.*
10: /cgn2_6/ptodata/2/paa/US086_COMB.pep.*
11: /cgn2_6/ptodata/2/paa/US087_COMB.pep.*
12: /cgn2_6/ptodata/2/paa/US088_COMB.pep.*
13: /cgn2_6/ptodata/2/paa/US089_COMB.pep.*
14: /cgn2_6/ptodata/2/paa/US090_COMB.pep.*
15: /cgn2_6/ptodata/2/paa/US091_COMB.pep.*
16: /cgn2_6/ptodata/2/paa/US092_COMB.pep.*
17: /cgn2_6/ptodata/2/paa/US093_COMB.pep.*
18: /cgn2_6/ptodata/2/paa/US094_COMB.pep.*
19: /cgn2_6/ptodata/2/paa/US095_COMB.pep.*
20: /cgn2_6/ptodata/2/paa/US096_COMB.pep.*
21: /cgn2_6/ptodata/2/paa/US097_COMB.pep.*
22: /cgn2_6/ptodata/2/paa/US098_COMB.pep.*
23: /cgn2_6/ptodata/2/paa/US60_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	480	100.0	480	US-08-480-229B-10	Sequence 10, Appl
2	480	100.0	480	US-08-659-235B-10	Sequence 10, Appl
3	480	100.0	480	US-09-237-981-10	Sequence 10, Appl
4	194	40.4	221	US-09-237-981-29	Sequence 29, Appl
5	79	16.5	513	US-08-480-229B-14	Sequence 14, Appl
6	79	16.5	513	US-08-659-235B-14	Sequence 14, Appl
7	79	16.5	513	US-09-237-981-14	Sequence 14, Appl
8	57	11.9	85	US-08-480-229B-1	Sequence 1, Appl
9	57	11.9	85	US-08-659-235B-1	Sequence 1, Appl
10	57	11.9	85	US-09-237-981-1	Sequence 1, Appl

11	51	10.6	311	23	US-60-230-435-1519	Sequence 1519, Ap
12	51	10.6	362	23	US-60-207-315-523	Sequence 523, App
13	51	10.6	449	23	US-60-207-315-404	Sequence 404, App
14	45	9.4	185	23	US-60-230-435-1806	Sequence 1806, Ap
15	45	9.4	203	23	US-60-233-644-77	Sequence 77, Appl
16	45	9.4	481	10	US-08-659-235B-29	Sequence 29, Appl
17	41	8.5	43	8	US-08-480-229B-24	Sequence 24, Appl
18	41	8.5	43	10	US-08-659-235B-24	Sequence 24, Appl
19	41	8.5	43	16	US-09-237-981-24	Sequence 24, Appl
20	40	8.3	52	23	US-60-236-359-21193	Sequence 21193, A
21	40	8.3	321	8	US-08-480-229B-21	Sequence 21, Appl
22	40	8.3	321	10	US-08-659-235B-21	Sequence 21, Appl
23	40	8.3	321	16	US-09-237-981-21	Sequence 21, Appl
24	38	7.9	48	23	US-60-236-359-20190	Sequence 20190, A
25	38	7.9	62	23	US-60-192-739-3075	Sequence 3075, Ap
26	36	7.5	50	23	US-60-171-481-1389	Sequence 1389, Ap
27	36	7.5	56	23	US-60-171-489-1052	Sequence 1052, Ap
28	30	6.2	36	23	US-60-160-189-7231	Sequence 7231, Ap
29	30	6.2	36	23	US-60-169-867-5120	Sequence 5120, Ap
30	27	5.6	29	23	US-60-160-203-4296	Sequence 4296, Ap
31	27	5.6	37	23	US-60-236-359-18942	Sequence 18942, A
32	27	5.6	54	23	US-60-170-346-1599	Sequence 1599, Ap
33	27	5.6	54	23	US-60-170-430-1864	Sequence 1864, Ap
34	24	5.0	43	23	US-60-160-189-6833	Sequence 6833, Ap
35	24	5.0	48	23	US-60-236-359-21021	Sequence 21021, A
36	21	4.4	33	23	US-60-192-739-4020	Sequence 4020, Ap
37	21	4.4	33	23	US-60-194-243-2893	Sequence 2893, Ap
38	21	4.4	42	23	US-60-169-840-5748	Sequence 5748, Ap
39	19	4.0	32	23	US-60-160-203-3897	Sequence 3897, Ap
40	18	3.8	25	8	US-08-480-229B-22	Sequence 22, Appl
41	18	3.8	25	10	US-08-659-235B-22	Sequence 22, Appl
42	18	3.8	25	16	US-09-237-981-22	Sequence 25, Appl
43	18	3.8	42	8	US-08-480-229B-25	Sequence 25, Appl
44	18	3.8	42	10	US-08-659-235B-25	Sequence 25, Appl
45	18	3.8	42	16	US-09-237-981-25	Sequence 25, Appl
46	17	3.5	34	23	US-60-177-646-2311	Sequence 2311, Ap
47	16	3.3	44	23	US-60-177-646-2310	Sequence 2310, Ap
48	14	2.9	57	8	US-08-480-229B-23	Sequence 23, Appl
49	14	2.9	57	10	US-08-659-235B-23	Sequence 23, Appl
50	14	2.9	57	16	US-09-237-981-23	Sequence 23, Appl

ALIGNMENTS

RESULT 1
US-08-480-229B-10
; Sequence 10, Application US/08480229B
; GENERAL INFORMATION:
; APPLICANT: Quentromous, Thomas
; APPLICANT: Hogan, Brigid
; APPLICANT: Snodgrass, H. Ralph
; APPLICANT: Zupancic, Thomas J.
; TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
; TITLE OF INVENTION: CELL LOCUS-1
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,229B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:

```
; NAME: Poissant, Brian M.
; REGISTRATION NUMBER: 28.462
; REFERENCE/DOCKET NUMBER: 8907-026
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 Pennie
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 480 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-480-229B-10

Query Match 100.0%; Score 480; DB 8; Length 480;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 480; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKHLVAWLLVGLSLGVPOFGKGDICNPNCENGICLSGLADDSFSCPEGFAGPNC 60
DB 1 MKHLVAWLLVGLSLGVPOFGKGDICNPNCENGICLSGLADDSFSCPEGFAGPNC 60
QY 61 SVVEVASDEEKPTSGAGPCIPNCHNGGTCEISEAYRGDTFYGVCPCPRGFGIHCQNI 120
DB 61 SVVEVASDEEKPTSGAGPCIPNCHNGGTCEISEAYRGDTFYGVCPCPRGFGIHCQNI 120
QY 121 NECEAPCRNGGICTDLVANYSCPCGPFMGRCQYKCSGHLGIEGGIISNQOITASSNH 180
DB 121 NECEAPCRNGGICTDLVANYSCPCGPFMGRCQYKCSGHLGIEGGIISNQOITASSNH 180
QY 181 RALFGLQKWPPYARLNKGLINAWTAENDRPWQIQLNQRKMRVTGVTQAKRIGSP 240
DB 181 RALFGLQKWPPYARLNKGLINAWTAENDRPWQIQLNQRKMRVTGVTQAKRIGSP 240
QY 241 EYIKSYKIAYSDGKTWAMYKVKGTNEEMVFRGNVDNTPYANSFTPIKQYVRLYPQI 300
DB 241 EYIKSYKIAYSDGKTWAMYKVKGTNEEMVFRGNVDNTPYANSFTPIKQYVRLYPQI 300
QY 301 CRRHCTLRMELLCGELSGCSEPLGMKSGHIGDYQITASSVFTLNMDFTWEPKARLDK 360
DB 301 CRRHCTLRMELLCGELSGCSEPLGMKSGHIGDYQITASSVFTLNMDFTWEPKARLDK 360
QY 361 QKVNANTSGHNDQSOWLQVLLVPTKVTGIIITQAKDFGHVQVGSYKLAYSDNGEHWM 420
DB 361 QKVNANTSGHNDQSOWLQVLLVPTKVTGIIITQAKDFGHVQVGSYKLAYSDNGEHWM 420
QY 421 VHODEKQKDKVFGQNFNDTHRKNVIDPPIYARFIRILPWSWYGRITLRSLLGCABEE 480
DB 421 VHODEKQKDKVFGQNFNDTHRKNVIDPPIYARFIRILPWSWYGRITLRSLLGCABEE 480

RESULT 2
US-08-659-235B-10
; GENERAL INFORMATION:
; APPLICANT: Quermous, Thomas
; APPLICANT: Hogan, Brigid
; APPLICANT: Snodgrass, H. Ralph
; APPLICANT: Zupancic, Thomas J.
; TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESS: Pennie & Edmonds LLP
; CITY: New York
; STATE: New York
; COUNTRY: United States
; *ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```

```
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/659,235B
; FILING DATE: 05-JUN-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
; REGISTRATION NUMBER: 28.462
; REFERENCE/DOCKET NUMBER: 8907-034
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 Pennie
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 480 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-659-235B-10

Query Match 100.0%; Score 480; DB 10; Length 480;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 480; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKHLVAWLLVGLSLGVPOFGKGDICNPNCENGICLSGLADDSFSCPEGFAGPNC 60
DB 1 MKHLVAWLLVGLSLGVPOFGKGDICNPNCENGICLSGLADDSFSCPEGFAGPNC 60
QY 61 SVVEVASDEEKPTSGAGPCIPNCHNGGTCEISEAYRGDTFYGVCPCPRGFGIHCQNI 120
DB 61 SVVEVASDEEKPTSGAGPCIPNCHNGGTCEISEAYRGDTFYGVCPCPRGFGIHCQNI 120
QY 121 NECEAPCRNGGICTDLVANYSCPCGPFMGRCQYKCSGHLGIEGGIISNQOITASSNH 180
DB 121 NECEAPCRNGGICTDLVANYSCPCGPFMGRCQYKCSGHLGIEGGIISNQOITASSNH 180
QY 181 RALFGLQKWPPYARLNKGLINAWTAENDRPWQIQLNQRKMRVTGVTQAKRIGSP 240
DB 181 RALFGLQKWPPYARLNKGLINAWTAENDRPWQIQLNQRKMRVTGVTQAKRIGSP 240
QY 241 EYIKSYKIAYSDGKTWAMYKVKGTNEEMVFRGNVDNTPYANSFTPIKQYVRLYPQI 300
DB 241 EYIKSYKIAYSDGKTWAMYKVKGTNEEMVFRGNVDNTPYANSFTPIKQYVRLYPQI 300
QY 301 CRRHCTLRMELLCGELSGCSEPLGMKSGHIGDYQITASSVFTLNMDFTWEPKARLDK 360
DB 301 CRRHCTLRMELLCGELSGCSEPLGMKSGHIGDYQITASSVFTLNMDFTWEPKARLDK 360
QY 361 QKVNANTSGHNDQSOWLQVLLVPTKVTGIIITQAKDFGHVQVGSYKLAYSDNGEHWM 420
DB 361 QKVNANTSGHNDQSOWLQVLLVPTKVTGIIITQAKDFGHVQVGSYKLAYSDNGEHWM 420
QY 421 VHODEKQKDKVFGQNFNDTHRKNVIDPPIYARFIRILPWSWYGRITLRSLLGCABEE 480
DB 421 VHODEKQKDKVFGQNFNDTHRKNVIDPPIYARFIRILPWSWYGRITLRSLLGCABEE 480

RESULT 3
US-09-237-981-10
; Sequence 10, Application US/09237981
; GENERAL INFORMATION:
; APPLICANT: Quermous, Thomas
; APPLICANT: Hogan, Brigid
; APPLICANT: Snodgrass, H. Ralph
; APPLICANT: Zupancic, Thomas J.
; TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
```


ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/237,981
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/659,235
FILING DATE: 05-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-0034-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennie
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 480 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-237-981-10

Query Match 100.0%; Score 480; DB 16; Length 480;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 480; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKHLVAALLVGLSLGVQFGKGDICNPENCGGICLSGLADDSFSCPEGFAGPNC 60
|||||
Db 1 MKHLVAALLVGLSLGVQFGKGDICNPENCGGICLSGLADDSFSCPEGFAGPNC 60
|||||

QY 61 SVVEVASDEEKPTSGAGPCIPNCHNGGTCISEAYRGDTFYGVCCKPRGFNGIHCOHNI 120
|||||
Db 61 SVVEVASDEEKPTSGAGPCIPNCHNGGTCISEAYRGDTFYGVCCKPRGFNGIHCOHNI 120
|||||

QY 121 NECEAEPCRNNGGICTDLVANSCECPGFMRNCQYKCSGHLGIEGGIISNOQITASSNH 180
|||||
Db 121 NECEAEPCRNNGGICTDLVANSCECPGFMRNCQYKCSGHLGIEGGIISNOQITASSNH 180
|||||

QY 181 RALFGLQKWYPYARLNKKGILINAWTAANDRWPMQINLQKMRVTGVTQGAKRIGSP 240
|||||
Db 181 RALFGLQKWYPYARLNKKGILINAWTAANDRWPMQINLQKMRVTGVTQGAKRIGSP 240
|||||

QY 241 EYIKSYKAYSNDGKTWAMYKVGTEEMVFRGNVDNNTPYANFTPPIKAQYVRLYPQI 300
|||||
Db 241 EYIKSYKAYSNDGKTWAMYKVGTEEMVFRGNVDNNTPYANFTPPIKAQYVRLYPQI 300
|||||

QY 301 CRRHCTLRWELGCGSCSEPLGKMSGHIODYQITASSVFTLNMDFTWEPKRLDK 360
|||||
Db 301 CRRHCTLRWELGCGSCSEPLGKMSGHIODYQITASSVFTLNMDFTWEPKRLDK 360
|||||

QY 361 QGKVNAWTSNGNDQSQWLQVLLVPTKVTGIIITQAKDFGHVQFVGSYKLAYSNDGEHWM 420
|||||
Db 361 QGKVNAWTSNGNDQSQWLQVLLVPTKVTGIIITQAKDFGHVQFVGSYKLAYSNDGEHWM 420
|||||

QY 421 VHODEKQRKDKVFQGNFNDTHRKNVIDPPIYARFIRLPWSWYGRITLRSELLGCAEE 480
|||||
Db 421 VHODEKQRKDKVFQGNFNDTHRKNVIDPPIYARFIRLPWSWYGRITLRSELLGCAEE 480
|||||

RESULT 4

US-09-237-981-29
Sequence 29, Application US/09237981
GENERAL INFORMATION:
APPLICANT: Quertemous, Thomas
APPLICANT: Hogan, Brigid
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
CELL LOCUS-1
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/237,981
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/659,235
FILING DATE: 05-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-0034-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennie
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 221 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-09-237-981-29

Query Match 40.4%; Score 194; DB 16; Length 221;
Best Local Similarity 100.0%; Pred. No. 1.3e-187;
Matches 194; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKHLVAALLVGLSLGVQFGKGDICNPENCGGICLSGLADDSFSCPEGFAGPNC 60
|||||
Db 1 MKHLVAALLVGLSLGVQFGKGDICNPENCGGICLSGLADDSFSCPEGFAGPNC 60
|||||

QY 61 SVVEVASDEEKPTSGAGPCIPNCHNGGTCISEAYRGDTFYGVCCKPRGFNGIHCOHNI 120
|||||
Db 61 SVVEVASDEEKPTSGAGPCIPNCHNGGTCISEAYRGDTFYGVCCKPRGFNGIHCOHNI 120
|||||

QY 121 NECEAEPCRNNGGICTDLVANSCECPGFMRNCQYKCSGHLGIEGGIISNOQITASSNH 180
|||||
Db 121 NECEAEPCRNNGGICTDLVANSCECPGFMRNCQYKCSGHLGIEGGIISNOQITASSNH 180
|||||

QY 181 RALFGLQKWYPY 194
|||||
Db 181 RALFGLQKWYPY 194
|||||

RESULT 5
US-08-480-229B-14
Sequence 14, Application US/08480229B
GENERAL INFORMATION:

APPLICANT: Quettermous, Thomas
APPLICANT: Hogan, Brigid
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,229B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-026
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennie
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 513 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-480-229B-14

Query Match 16.5% Score 79; DB 8; Length 513;
Best Local Similarity 100.0%; Pred. No. 5.7e-71;
Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 341 FRTLNDMFTWEPKARLDKQGVNWTSGHNDQSQWLQVDLLVPTKVTGIIITQAKDFG 400
|||||
Db 374 FRTLNDMFTWEPKARLDKQGVNWTSGHNDQSQWLQVDLLVPTKVTGIIITQAKDFG 433
|||||

Qy 401 HVQFVGSYKLAYSNDGEHW 419
|||||
Db 434 HVQFVGSYKLAYSNDGEHW 452
|||||

RESULT 6
US-08-659-235B-14
Sequence 14, Application US/08659235B
GENERAL INFORMATION:
APPLICANT: Quettermous, Thomas
APPLICANT: Hogan, Brigid
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/659,235B
FILING DATE: 05-JUN-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-034
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennie
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 513 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-659-235B-14

Query Match 16.5% Score 79; DB 10; Length 513;
Best Local Similarity 100.0%; Pred. No. 5.7e-71;
Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 341 FRTLNDMFTWEPKARLDKQGVNWTSGHNDQSQWLQVDLLVPTKVTGIIITQAKDFG 400
|||||
Db 374 FRTLNDMFTWEPKARLDKQGVNWTSGHNDQSQWLQVDLLVPTKVTGIIITQAKDFG 433
|||||

Qy 401 HVQFVGSYKLAYSNDGEHW 419
|||||
Db 434 HVQFVGSYKLAYSNDGEHW 452
|||||

RESULT 7
US-09-237-981-14
Sequence 14, Application US/09237981
GENERAL INFORMATION:
APPLICANT: Quettermous, Thomas
APPLICANT: Hogan, Brigid
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/237,981
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/659,235
FILING DATE: 05-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-0034-999

TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennie
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 513 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-237-981-14

Query Match 16.5%; Score 79; DB 16; Length 513;
Best Local Similarity 100.0%; Pred. No. 5.7e-71;
Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 341 FRTLNDMTWPRKARLDKQGVNWTSGHNDQSQWLQVLLVPTKVTGIITQAKDFG 400
|||||
Db 374 FRTLNDMTWPRKARLDKQGVNWTSGHNDQSQWLQVLLVPTKVTGIITQAKDFG 433
QY 401 HVQFVGSYKLAYSNDGEHW 419
Db 434 HVQFVGSYKLAYSNDGEHW 452

RESULT 8
US-08-480-229B-1
; Sequence 1, Application US/08480229B
; GENERAL INFORMATION:
; APPLICANT: Quertermous, Thomas
; APPLICANT: Hogan, Brigid
; APPLICANT: Snodgrass, H. Ralph
; APPLICANT: Zupancic, Thomas J.
; TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2711
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,229B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-026
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 Pennie
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 85 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-480-229B-1

Query Match 11.9%; Score 57; DB 8; Length 85;

Best Local Similarity 100.0%; Pred. No. 2e-49;
Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 404 FVGSYKLAYSNDGEHWMVHQDEKQKDKVFGQNFNDTHRKNVIDPPIYARFIRILP 460
|||||
Db 26 FVGSYKLAYSNDGEHWMVHQDEKQKDKVFGQNFNDTHRKNVIDPPIYARFIRILP 82

RESULT 9
US-08-659-235B-1
; Sequence 1, Application US/08659235B
; GENERAL INFORMATION:
; APPLICANT: Quertermous, Thomas
; APPLICANT: Hogan, Brigid
; APPLICANT: Snodgrass, H. Ralph
; APPLICANT: Zupancic, Thomas J.
; TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2711
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/659,235B
; FILING DATE: 05-JUN-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-034
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 Pennie
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 85 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-659-235B-1

Query Match 11.9%; Score 57; DB 10; Length 85;
Best Local Similarity 100.0%; Pred. No. 2e-49;
Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 404 FVGSYKLAYSNDGEHWMVHQDEKQKDKVFGQNFNDTHRKNVIDPPIYARFIRILP 460
|||||
Db 26 FVGSYKLAYSNDGEHWMVHQDEKQKDKVFGQNFNDTHRKNVIDPPIYARFIRILP 82

RESULT 10
US-09-237-981-1
; Sequence 1, Application US/09237981
; GENERAL INFORMATION:
; APPLICANT: Quertermous, Thomas
; APPLICANT: Hogan, Brigid
; APPLICANT: Snodgrass, H. Ralph
; APPLICANT: Zupancic, Thomas J.
; TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
; NUMBER OF SEQUENCES: 29

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Pennie & Edmonds LLP
;; STREET: 1155 Avenue of the Americas
;; CITY: New York
;; STATE: New York
;; COUNTRY: United States
;; ZIP: 10036-2711
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/237,981
;; FILING DATE:
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/08/659,235
;; FILING DATE: 05-JUN-1996
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Poissant, Brian M.
;; REGISTRATION NUMBER: 28,462
;; REFERENCE/DOCKET NUMBER: 8907-0034-999
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (212) 790-3090
;; TELEFAX: (212) 869-8884/9741
;; TELEX: 66141 Pennie
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 85 amino acids
;; TYPE: amino acid
;; STRANDEDNESS:
;; TOPOLOGY: unknown
;; MOLECULE TYPE: peptide
US-09-237-981-1

Query Match 11.9%; Score 57; DB 16; Length 85;
Best Local Similarity 100.0%; Pred. No. 2e-49;
Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 404 FVGSYKLAYSNDGEHVMVHODEKQKDKVFGNFDNDRKKNVIDPPIYARFIRILP 460
|||||
Db 26 FVGSYKLAYSNDGEHVMVHODEKQKDKVFGNFDNDRKKNVIDPPIYARFIRILP 82

RESULT 11
US-60-230-435-1519
; Sequence 1519, Application US/60230435
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL000768
; CURRENT APPLICATION NUMBER: US/60/230,435
; CURRENT FILING DATE: 2000-09-06
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1519
; LENGTH: 311
; TYPE: PRT
; ORGANISM: HUMAN
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(311)
; OTHER INFORMATION: Xaa - Any Amino Acid
US-60-230-435-1519

Query Match 10.6%; Score 51; DB 23; Length 311;
Best Local Similarity 100.0%; Pred. No. 7.7e-43;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 217 QINLQKRMVTVITQGAKRIGSPYIKSYKIAYSNDGKTWAMYKVKGTNE 267
|||||
Db 111 QINLQKRMVTVITQGAKRIGSPYIKSYKIAYSNDGKTWAMYKVKGTNE 161

RESULT 12
US-60-207-315-523
; Sequence 523, Application US/60207315
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL000601
; CURRENT APPLICATION NUMBER: US/60/207,315
; CURRENT FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 528
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 523
; LENGTH: 362
; TYPE: PRT
; ORGANISM: HUMAN
US-60-207-315-523

Query Match 10.6%; Score 51; DB 23; Length 362;
Best Local Similarity 100.0%; Pred. No. 8.8e-43;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 217 QINLQKRMVTVITQGAKRIGSPYIKSYKIAYSNDGKTWAMYKVKGTNE 267
|||||
Db 162 QINLQKRMVTVITQGAKRIGSPYIKSYKIAYSNDGKTWAMYKVKGTNE 212

RESULT 13
US-60-207-315-404
; Sequence 404, Application US/60207315
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL000601
; CURRENT APPLICATION NUMBER: US/60/207,315
; CURRENT FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 528
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 404
; LENGTH: 449
; TYPE: PRT
; ORGANISM: HUMAN
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(449)
; OTHER INFORMATION: Xaa - Any Amino Acid
US-60-207-315-404

Query Match 10.6%; Score 51; DB 23; Length 449;
Best Local Similarity 100.0%; Pred. No. 1.1e-42;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 217 QINLQKRMVTVITQGAKRIGSPYIKSYKIAYSNDGKTWAMYKVKGTNE 267
|||||
Db 249 QINLQKRMVTVITQGAKRIGSPYIKSYKIAYSNDGKTWAMYKVKGTNE 299

RESULT 14
US-60-230-435-1806
; Sequence 1806, Application US/60230435
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen

; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL000768
; CURRENT APPLICATION NUMBER: US/60/230,435
; CURRENT FILING DATE: 2000-09-06
; NUMBER OF SEQ ID NOS: 2991
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1806
; LENGTH: 185
; TYPE: PRT
; ORGANISM: HUMAN
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)---(185)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-60-230-435-1806

Query Match 9.4%; Score 45; DB 23; Length 185;
Best Local Similarity 100.0%; Pred. No. 5.6e-37;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 PNPCHNGTCEISEAYRGDTFGYVCKCPRGFNGIHCHQHNINECE 124
|||||
Db 48 PNPCHNGTCEISEAYRGDTFGYVCKCPRGFNGIHCHQHNINECE 92

RESULT 15
US-60-233-644-77
; Sequence 77, Application US/60233644
; GENERAL INFORMATION:

; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN NUCLEAR HORMONE RECEPTOR
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN NUCLEAR
; TITLE OF INVENTION: HORMONE RECEPTOR PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000824
; CURRENT APPLICATION NUMBER: US/60/233,644
; CURRENT FILING DATE: 2000-09-18
; NUMBER OF SEQ ID NOS: 150
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 77
; LENGTH: 203
; TYPE: PRT
; ORGANISM: HUMAN
US-60-233-644-77

Query Match 9.4%; Score 45; DB 23; Length 203;
Best Local Similarity 100.0%; Pred. No. 6.1e-37;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 PNPCHNGTCEISEAYRGDTFGYVCKCPRGFNGIHCHQHNINECE 124
|||||
Db 58 PNPCHNGTCEISEAYRGDTFGYVCKCPRGFNGIHCHQHNINECE 102

RESULT 16
US-08-659-235B-29
; Sequence 29, Application US/08659235B
; GENERAL INFORMATION:

; APPLICANT: Quattermous, Thomas
; APPLICANT: Hogan, Brigid
; APPLICANT: Snodgrass, H. Ralph
; APPLICANT: Zupancic, Thomas J.
; TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
; TITLE OF INVENTION: CELL LOCUS-1
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York

; COUNTRY: United States
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/659,235B
; FILING DATE: 05-JUN-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-034
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 Pennie
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 481 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-659-235B-29

Query Match 9.4%; Score 45; DB 10; Length 481;
Best Local Similarity 100.0%; Pred. No. 1.3e-36;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 PNPCHNGTCEISEAYRGDTFGYVCKCPRGFNGIHCHQHNINECE 124
|||||
Db 80 PNPCHNGTCEISEAYRGDTFGYVCKCPRGFNGIHCHQHNINECE 124

RESULT 17
US-08-480-229B-24
; Sequence 24, Application US/08480229B
; GENERAL INFORMATION:
; APPLICANT: Quattermous, Thomas
; APPLICANT: Hogan, Brigid
; APPLICANT: Snodgrass, H. Ralph
; APPLICANT: Zupancic, Thomas J.
; TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
; TITLE OF INVENTION: CELL LOCUS-1
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,229B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-026
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 Pennie
; INFORMATION FOR SEQ ID NO: 24:

;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 43 amino acids
;; TYPE: amino acid
;; STRANDEDNESS:
;; TOPOLOGY: unknown
;; MOLECULE TYPE: peptide
US-08-480-229B-24

Query Match 8.5%; Score 41; DB 8; Length 43;
Best Local Similarity 100.0%; Pred. No. 1.7e-33;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 PNPCHNGTCEISEAYRGDTFTGYVCKPRGFGNGIHCQHNI 120
|||||
DB 3 PNPCHNGTCEISEAYRGDTFTGYVCKPRGFGNGIHCQHNI 43

RESULT 18

US-08-659-235B-24
;; Sequence 24, Application US/08659235B
;; GENERAL INFORMATION:
;; APPLICANT: Quertermous, Thomas
;; APPLICANT: Hogan, Brigid
;; APPLICANT: Snodgrass, H. Ralph
;; APPLICANT: Zupancic, Thomas J.
;; TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
;; TITLE OF INVENTION: CELL LOCUS-1
;; NUMBER OF SEQUENCES: 29
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Pennie & Edmonds LLP
;; STREET: 1155 Avenue of the Americas
;; CITY: New York
;; STATE: New York
;; COUNTRY: United States
;; ZIP: 10036-2711
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/659,235B
;; FILING DATE: 05-JUN-1996
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Poissant, Brian M.
;; REGISTRATION NUMBER: 28,462
;; REFERENCE/DOCKET NUMBER: 8907-034
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (212) 790-9090
;; TELEFAX: (212) 869-8864/9741
;; TELEX: 66141 Pennie
;; INFORMATION FOR SEQ ID NO: 24:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 43 amino acids
;; TYPE: amino acid
;; STRANDEDNESS:
;; TOPOLOGY: unknown
;; MOLECULE TYPE: peptide
US-08-659-235B-24

Query Match 8.5%; Score 41; DB 10; Length 43;
Best Local Similarity 100.0%; Pred. No. 1.7e-33;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 PNPCHNGTCEISEAYRGDTFTGYVCKPRGFGNGIHCQHNI 120
|||||
DB 3 PNPCHNGTCEISEAYRGDTFTGYVCKPRGFGNGIHCQHNI 43

RESULT 19

;; SOFTWARE: Molecular Dynamics Sequence Listing Engine

US-09-237-981-24
;; Sequence 24, Application US/09237981
;; GENERAL INFORMATION:
;; APPLICANT: Quertermous, Thomas
;; APPLICANT: Hogan, Brigid
;; APPLICANT: Snodgrass, H. Ralph
;; APPLICANT: Zupancic, Thomas J.
;; TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
;; TITLE OF INVENTION: CELL LOCUS-1
;; NUMBER OF SEQUENCES: 29
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Pennie & Edmonds LLP
;; STREET: 1155 Avenue of the Americas
;; CITY: New York
;; STATE: New York
;; COUNTRY: United States
;; ZIP: 10036-2711
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/237,981
;; FILING DATE:
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/08/659,235
;; FILING DATE: 05-JUN-1996
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Poissant, Brian M.
;; REGISTRATION NUMBER: 28,462
;; REFERENCE/DOCKET NUMBER: 8907-0034-999
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (212) 790-9090
;; TELEFAX: (212) 869-8864/9741
;; TELEX: 66141 Pennie
;; INFORMATION FOR SEQ ID NO: 24:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 43 amino acids
;; TYPE: amino acid
;; STRANDEDNESS:
;; TOPOLOGY: unknown
;; MOLECULE TYPE: peptide
US-09-237-981-24

Query Match 8.5%; Score 41; DB 16; Length 43;
Best Local Similarity 100.0%; Pred. No. 1.7e-33;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 PNPCHNGTCEISEAYRGDTFTGYVCKPRGFGNGIHCQHNI 120
|||||
DB 3 PNPCHNGTCEISEAYRGDTFTGYVCKPRGFGNGIHCQHNI 43

RESULT 20

US-60-236-359-21193
;; Sequence 21193, Application US/60236359
;; GENERAL INFORMATION:
;; APPLICANT: Penn, Sharon G.
;; APPLICANT: Rank, David R.
;; APPLICANT: Hanzel, David K.
;; APPLICANT: Chen, Wensheng
;; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL F
;; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
;; FILE REFERENCE: MDHMORF-4P
;; CURRENT APPLICATION NUMBER: US/60/236,359
;; CURRENT FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: US 60/207,456
;; PRIOR FILING DATE: 2000-05-26
;; NUMBER OF SEQ ID NOS: 21709
;; SOFTWARE: Molecular Dynamics Sequence Listing Engine

SEQ ID NO 21193
LENGTH: 52

TYPE: PRT
ORGANISM: Homo sapiens

FEATURE:

OTHER INFORMATION: MAP TO AC008430.2

OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1

OTHER INFORMATION: SWISSPROT HIT: P70490, EVALUE 3.00e-13

OTHER INFORMATION: EST_HUMAN HIT: AW965338.1, EVALUE 6.00e-25

US-60-236-359-21193

Query Match 8.3%; Score 40; DB 23; Length 52;
Best Local Similarity 100.0%; Pred. No. 2e-32;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 380 VLLVPTKVTGIIQAKDFGHVQFVGSYKLYSNDGEHW 419
|||||
DB 1 VLLVPTKVTGIIQAKDFGHVQFVGSYKLYSNDGEHW 40

RESULT 21

US-08-480-229B-21

Sequence 21, Application US/08480229B

GENERAL INFORMATION:

APPLICANT: Quertermous, Thomas

APPLICANT: Hogan, Brigid

APPLICANT: Snodgrass, H. Ralph

APPLICANT: Zupancic, Thomas J.

TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL

TITLE OF INVENTION: CELL LOCUS-1

NUMBER OF SEQUENCES: 28

CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: United States

ZIP: 10036-2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/480,229B

FILING DATE: 07-JUN-1995

CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:

NAME: Poissant, Brian M.

REGISTRATION NUMBER: 28,462

REFERENCE/DOCKET NUMBER: 8907-026

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 790-9090

TELEFAX: (212) 869-8864/9741

TELEX: 66141 Pennie

INFORMATION FOR SEQ ID NO: 21:

SEQUENCE CHARACTERISTICS:

LENGTH: 321 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: unknown

MOLECULE TYPE: protein

US-08-480-229B-21

Query Match 8.3%; Score 40; DB 8; Length 321;
Best Local Similarity 100.0%; Pred. No. 1.1e-31;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 341 FRTLNDMFTWEPKRLDKQKYNWNTSGHNDQSOWLQV 380
|||||
DB 185 FRTLNDMFTWEPKRLDKQKYNWNTSGHNDQSOWLQV 224

RESULT 22

US-08-659-235B-21

Sequence 21, Application US/08659235B

GENERAL INFORMATION:

APPLICANT: Quertermous, Thomas

APPLICANT: Hogan, Brigid

APPLICANT: Snodgrass, H. Ralph

APPLICANT: Zupancic, Thomas J.

TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL

TITLE OF INVENTION: CELL LOCUS-1

NUMBER OF SEQUENCES: 29

CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds LLP

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: United States

ZIP: 10036-2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/659,235B

FILING DATE: 05-JUN-1996

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Poissant, Brian M.

REGISTRATION NUMBER: 28,462

REFERENCE/DOCKET NUMBER: 8907-034

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 790-9090

TELEFAX: (212) 869-8864/9741

INFORMATION FOR SEQ ID NO: 21:

SEQUENCE CHARACTERISTICS:

LENGTH: 321 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: unknown

MOLECULE TYPE: protein

US-08-659-235B-21

Query Match 8.3%; Score 40; DB 10; Length 321;
Best Local Similarity 100.0%; Pred. No. 1.1e-31;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 341 FRTLNDMFTWEPKRLDKQKYNWNTSGHNDQSOWLQV 380
|||||
DB 185 FRTLNDMFTWEPKRLDKQKYNWNTSGHNDQSOWLQV 224

RESULT 23

US-09-237-981-21

Sequence 21, Application US/09237981

GENERAL INFORMATION:

APPLICANT: Quertermous, Thomas

APPLICANT: Hogan, Brigid

APPLICANT: Snodgrass, H. Ralph

APPLICANT: Zupancic, Thomas J.

TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL

TITLE OF INVENTION: CELL LOCUS-1

NUMBER OF SEQUENCES: 29

CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds LLP

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: United States

ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
FILING DATE: US/09/237,981
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/659,235
FILING DATE: 05-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-0034-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennie
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 321 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-09-237-981-21

Query Match 8.3%; Score 40; DB 16; Length 321;
Best Local Similarity 100.0%; Pred. No. 1.1e-31;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 341 FRTLNDMFTWEPKARLDKQKGNWNTSGHNDQSQWLQV 380
Db 185 FRTLNDMFTWEPKARLDKQKGNWNTSGHNDQSQWLQV 224

RESULT 24

US-60-236-359-20190
Sequence 20190, Application US/60236359
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: MDHMOF-4P
CURRENT APPLICATION NUMBER: US/60/236,359
CURRENT FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
NUMBER OF SEQ ID NOS: 21709
SOFTWARE: Molecular Dynamics Sequence Listing Engine
SEQ ID NO 20190
LENGTH: 48
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC008430.2
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.5
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.97
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.92
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.9
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 0.99
OTHER INFORMATION: EST_HUMAN HIT: RA112613.1, EVALUE 1.00e-19
OTHER INFORMATION: SWISSPROT HIT: P79385, EVALUE 2.00e-10

US-60-236-359-20190

Query Match 7.9%; Score 38; DB 23; Length 48;
Best Local Similarity 100.0%; Pred. No. 2e-30;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 180 HRALFGLQKWYPYARLNKKGLINAWTAENDRWPWQ 217
Db 11 HRALFGLQKWYPYARLNKKGLINAWTAENDRWPWQ 48

RESULT 25

US-60-192-739-3075
Sequence 3075, Application US/60192739
GENERAL INFORMATION:
APPLICANT: Bonazzi, Vivien
TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS,
TITLE OF INVENTION: USES THEREOF
FILE REFERENCE: CL000406
CURRENT APPLICATION NUMBER: US/60/192,739
CURRENT FILING DATE: 2000-03-28
NUMBER OF SEQ ID NOS: 4532
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3075
LENGTH: 62
TYPE: PRT
ORGANISM: HUMAN
US-60-192-739-3075

Query Match 7.9%; Score 38; DB 23; Length 62;
Best Local Similarity 100.0%; Pred. No. 2.5e-30;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 180 HRALFGLQKWYPYARLNKKGLINAWTAENDRWPWQ 217
Db 24 HRALFGLQKWYPYARLNKKGLINAWTAENDRWPWQ 61

Search completed: May 23, 2001, 06:23:34
Job time: 268 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 06:20:56 ; Search time 9.39 Seconds
(without alignments)

366.422 Million cell updates/sec

Title: US-09-237-981-10
Perfect score: 480
Sequence: 1 MKHLVAALLVGLGVQF.....WSWYGRITLRSELLGCAEE 480

Scoring table: OLIGO

Gapop 60.0 , Gapext 50.0

Searched: 41833 seqs, 7168127 residues

Word size : 0

Total number of hits satisfying chosen parameters: 41833

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 50 summaries

Database : Pending_Patents_AA_New.*

- 1: /cgn2_6/ptodata/1/paa/PCT_NEW_COMB.pep.*
- 2: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pep.*
- 3: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep.*
- 4: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep.*
- 5: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep.*
- 6: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	9	1.9	923	5	US-09-583-638-2
2	9	1.9	957	1	PCT-US00-35017A-1285
3	7	1.5	383	1	PCT-US01-11988-770
4	7	1.5	383	5	US-09-833-245-770
5	7	1.5	399	1	PCT-US01-04098A-1480
6	7	1.5	527	5	US-09-422-601-1
7	7	1.5	750	6	US-60-248-505-1003
8	7	1.5	795	6	US-60-248-505-1081
9	7	1.5	1697	1	PCT-US01-04098A-1861
10	7	1.5	4545	5	US-09-750-972-2
11	6	1.2	14	1	PCT-US01-06436-36
12	6	1.2	14	1	PCT-US01-06436-39
13	6	1.2	66	1	PCT-US01-01238-42
14	6	1.2	81	5	US-09-739-449-10947
15	6	1.2	83	5	US-09-739-449-1088
16	6	1.2	86	5	US-09-640-211A-805
17	6	1.2	100	5	US-09-739-449-11762
18	6	1.2	130	1	PCT-US00-35017A-773
19	6	1.2	138	1	PCT-US01-01339-4492
20	6	1.2	138	1	PCT-US01-01329-1749
21	6	1.2	145	5	US-09-814-666-274
22	6	1.2	148	1	PCT-US01-01302-91
23	6	1.2	186	1	PCT-US01-11988-924
24	6	1.2	186	5	US-09-833-245-924
25	6	1.2	188	1	PCT-US01-01302-109
26	6	1.2	197	1	PCT-US01-11988-1993
27	6	1.2	197	5	US-09-833-245-1993

28	6	1.2	205	4	US-08-706-945B-135	Sequence 135, App
29	6	1.2	205	5	US-09-811-284-154	Sequence 154, App
30	6	1.2	223	5	US-09-739-449-9657	Sequence 9657, App
31	6	1.2	229	5	US-09-731-924A-2	Sequence 2, Appli
32	6	1.2	234	5	US-09-640-211A-614	Sequence 614, App
33	6	1.2	243	5	US-09-739-449-11421	Sequence 11421, A
34	6	1.2	244	5	US-09-308-823A-596	Sequence 596, App
35	6	1.2	245	1	PCT-US01-11988-922	Sequence 922, App
36	6	1.2	245	5	US-09-833-245-922	Sequence 922, App
37	6	1.2	257	5	US-09-754-368-2	Sequence 2, Appli
38	6	1.2	259	1	PCT-US01-04098A-1328	Sequence 1328, Ap
39	6	1.2	279	1	PCT-US01-04098A-3296	Sequence 3296, Ap
40	6	1.2	283	4	US-08-956-171C-5203	Sequence 5203, Ap
41	6	1.2	288	1	PCT-US01-01302-107	Sequence 107, App
42	6	1.2	298	5	US-09-739-449-8950	Sequence 8950, Ap
43	6	1.2	304	5	US-60-282-814-5	Sequence 11600, A
44	6	1.2	327	6	US-09-202-059A-3	Sequence 5, Appli
45	6	1.2	349	5	US-09-739-449-9639	Sequence 9639, Ap
46	6	1.2	359	5	US-09-739-449-12197	Sequence 12197, A
47	6	1.2	365	5	US-09-808-689-4	Sequence 4, Appli
48	6	1.2	379	5	US-09-423-844-4	Sequence 4, Appli
49	6	1.2	379	5	US-09-403-296A-4	Sequence 4, Appli
50	6	1.2	379	5		

ALIGNMENTS

RESULT 1
US-09-583-638-2
; Sequence 2, Application US/09583638
; GENERAL INFORMATION:
; APPLICANT: KLAGSBRUN, MICHAEL
; APPLICANT: SOKER, SHAY
; APPLICANT: MIAO, HUA-QUAN
; APPLICANT: TAKASHIMA, SEIJI
; TITLE OF INVENTION: NEUROFILINS AND USE THEREOF IN METHODS FOR
; FILE OF INVENTION: DIAGNOSIS AND PROGNOSIS OF CANCER
; FILE REFERENCE: 701039-48800
; CURRENT APPLICATION NUMBER: US/09/583,638
; CURRENT FILING DATE: 2000-05-30
; PRIOR APPLICATION NUMBER: PCT/US98/26127
; PRIOR FILING DATE: 1998-12-09
; PRIOR APPLICATION NUMBER: 60/069,155
; PRIOR FILING DATE: 1997-12-09
; PRIOR APPLICATION NUMBER: 60/069,687
; PRIOR FILING DATE: 1997-12-12
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 923
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-583-638-2

Query Match 1.9%; Score 9; DB 5; Length 923;
Best Local Similarity 100.0%; Pred.No. 0.16;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 307 LRMELIGCE 315
Db 576 LRMELIGCE 584

RESULT 2
PCT-US00-35017A-1285
; Sequence 1285, Application PC/TUS0035017A
; GENERAL INFORMATION:
; APPLICANT: Hyseq Inc
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 784PCT
; CURRENT APPLICATION NUMBER: PCT/US00/35017A

; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US09/552,317
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 1478
; SEQ ID NO 1285
; LENGTH: 957
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US00-35017A-1285

Query Match 1.9%; Score 9; DB 1; Length 957;
Best Local Similarity 100.0%; Pred. No. 0.16; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0;

QY 307 LRMELLCGE 315
Db 610 LRMELLCGE 618
|||||

RESULT 3
PCT-US01-11988-770
; Sequence 770, Application PC/TUS0111988
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546PCT
; CURRENT APPLICATION NUMBER: PCT/US01/11988
; CURRENT FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: 60/229, 358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256, 931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199, 384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 770
; LENGTH: 383
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-11988-770

Query Match 1.5%; Score 7; DB 1; Length 383;
Best Local Similarity 100.0%; Pred. No. 9.1; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0;

QY 50 CPEGFAG 56
Db 200 CPEGFAG 206
|||||

RESULT 4
US-09-833-245-770
; Sequence 770, Application US/09833245
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546PCT
; CURRENT APPLICATION NUMBER: US/09/833,245
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229, 358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256, 931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199, 384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 770

; LENGTH: 383
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-833-245-770

Query Match 1.5%; Score 7; DB 5; Length 383;
Best Local Similarity 100.0%; Pred. No. 9.1; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0;

QY 50 CPEGFAG 56
Db 200 CPEGFAG 206
|||||

RESULT 5
PCT-US01-04098A-1480
; Sequence 1480, Application PC/TUS0104098A
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-029
; CURRENT APPLICATION NUMBER: PCT/US01/04098A
; CURRENT FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: Not Yet Assigned
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 09/728,422
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 09/693,325
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/663,561
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 09/654,936
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 09/620,325
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/598,075
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 3960
; SOFTWARE: Custom
; SEQ ID NO 1480
; LENGTH: 399
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-04098A-1480

Query Match 1.5%; Score 7; DB 1; Length 399;
Best Local Similarity 100.0%; Pred. No. 9.4; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0;

QY 381 DLLVPTK 387
Db 385 DLLVPTK 391
|||||

RESULT 6
US-09-422-601-1
; Sequence 1, Application US/09422601
; GENERAL INFORMATION:
; APPLICANT: Bajzar, Laszlo S.
; APPLICANT: Nesheim, Michael E.
; APPLICANT: Church, William R.
; TITLE OF INVENTION: Compositions for Inhibiting the
; TITLE OF INVENTION: Activation of Thrombin-Activatable Fibrinolysis Inhibitor
; TITLE OF INVENTION: (TAFI)
; FILE REFERENCE: VO139/7036 (HCL/MAT)
; CURRENT APPLICATION NUMBER: US/09/422,601
; CURRENT FILING DATE: 1999-10-21

;; PRIOR APPLICATION NUMBER: US 08/966,432
;; PRIOR FILING DATE: 1997-11-07
;; PRIOR APPLICATION NUMBER: US 60/030,721
;; PRIOR FILING DATE: 1996-11-08
;; NUMBER OF SEQ ID NOS: 1
;; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 527
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-422-601-1

Query Match 1.5%; Score 7; DB 5; Length 527;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 CPEGFAG 56
Db 75 CPEGFAG 81
|||||

RESULT 7
US-60-248-505-1003
; Sequence 1003, Application US/60248505
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
; FILE REFERENCE: c1000918
; CURRENT APPLICATION NUMBER: US/60/248,505
; CURRENT FILING DATE: 2000-11-15
; NUMBER OF SEQ ID NOS: 1998
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1003
; LENGTH: 750
; TYPE: PRT
; ORGANISM: Human
US-60-248-505-1003

Query Match 1.5%; Score 7; DB 6; Length 750;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 WLLVGLS 14
Db 411 WLLVGLS 417
|||||

RESULT 8
US-60-248-505-1081
; Sequence 1081, Application US/60248505
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
; FILE REFERENCE: c1000918
; CURRENT APPLICATION NUMBER: US/60/248,505
; CURRENT FILING DATE: 2000-11-15
; NUMBER OF SEQ ID NOS: 1998
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1081
; LENGTH: 795
; TYPE: PRT
; ORGANISM: Human
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(795)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-60-248-505-1081

Query Match 1.5%; Score 7; DB 6; Length 795;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 WLLVGLS 14
Db 416 WLLVGLS 422
|||||

RESULT 9
PCT-US01-04098A-1861
; Sequence 1861, Application PC/TUS0104098A
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-029
; CURRENT APPLICATION NUMBER: PCT/US01/04098A
; CURRENT FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: Not Yet Assigned
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 09/728,422
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 09/693,325
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/663,561
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 09/654,936
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 09/620,325
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/598,075
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 3960
; SOFTWARE: Custom
; SEQ ID NO 1861
; LENGTH: 1697
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-04098A-1861

Query Match 1.5%; Score 7; DB 1; Length 1697;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 425 EKQRKDK 431
Db 737 EKQRKDK 743
|||||

RESULT 10
US-09-750-972-2
; Sequence 2, Application US/09750972
; GENERAL INFORMATION:
; APPLICANT: Pramod K. Srivastava
; TITLE OF INVENTION: ALPHA(2) MACROGLOBULIN RECEPTOR AS A HEAT SHOCK
; TITLE OF INVENTION: PROTEIN RECEPTOR AND USES THEREOF
; FILE REFERENCE: 8449-134
; CURRENT APPLICATION NUMBER: US/09/750,972
; CURRENT FILING DATE: 2000-12-28
; PRIOR APPLICATION NUMBER: 09/750,972
; PRIOR FILING DATE: 2000-12-28
; PRIOR APPLICATION NUMBER: 09/668,724
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2

; LENGTH: 4545
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-750-972-2

Query Match 1.2%; Score 7; DB 5; Length 4545;
Best Local Similarity 100.0%; Pred. No. 84;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 83 CHNGGTC 89
|||||||
Db 4241 CHNGGTC 4247

RESULT 11
PCT-US01-06436-36
; Sequence 36, Application PC/TUS0106436
; GENERAL INFORMATION:

; APPLICANT: Hartnett, James R.

; APPLICANT: Huang, Fen

; TITLE OF INVENTION: Thermophilic DNA Polymerases from Thermoactinomyces

; FILE REFERENCE: PRMG-03819

; CURRENT APPLICATION NUMBER: PCT/US01/06436
; CURRENT FILING DATE: 2001-02-28

; NUMBER OF SEQ ID NOS: 39

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 36

; LENGTH: 14

; TYPE: PRT

; ORGANISM: T. maritima

PCT-US01-06436-36

Query Match 1.2%; Score 6; DB 1; Length 14;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 139 ANYSCE 144
|||||||
Db 4 ANYSCE 9

RESULT 12
PCT-US01-06436-39
; Sequence 39, Application PC/TUS0106436
; GENERAL INFORMATION:

; APPLICANT: Hartnett, James R.

; APPLICANT: Huang, Fen

; TITLE OF INVENTION: Thermophilic DNA Polymerases from Thermoactinomyces

; FILE REFERENCE: PRMG-03819

; CURRENT APPLICATION NUMBER: PCT/US01/06436
; CURRENT FILING DATE: 2001-02-28

; NUMBER OF SEQ ID NOS: 39

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 39

; LENGTH: 14

; TYPE: PRT

; ORGANISM: T. neapolitana

PCT-US01-06436-39

Query Match 1.2%; Score 6; DB 1; Length 14;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 139 ANYSCE 144
|||||||
Db 4 ANYSCE 9

RESULT 13

PCT-US01-01238-42

; Sequence 42, Application PC/TUS0101238
; GENERAL INFORMATION:

; APPLICANT: Human Genome Sciences, Inc., et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

; FILE REFERENCE: PA122PCT

; CURRENT APPLICATION NUMBER: PCT/US01/01238

; CURRENT FILING DATE: 2001-01-17

; NUMBER OF SEQ ID NOS: 137

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 42

; LENGTH: 66

; TYPE: PRT

; ORGANISM: Homo sapiens

PCT-US01-01238-42

Query Match 1.2%; Score 6; DB 1; Length 66;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 158 CSGHLG 163
|||||||
Db 4 CSGHLG 9

RESULT 14

US-09-739-449-10947

; Sequence 10947, Application US/09739449
; GENERAL INFORMATION:

; APPLICANT: Hinkle, Gregory J.

; APPLICANT: Slater, Steven C.

; TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof

; FILE REFERENCE: 38-10(15490)C

; CURRENT APPLICATION NUMBER: US/09/739,449

; CURRENT FILING DATE: 2000-12-19

; PRIOR APPLICATION NUMBER: US 09/514,000

; PRIOR FILING DATE: 2000-02-23

; NUMBER OF SEQ ID NOS: 13351

; SEQ ID NO 10947

; LENGTH: 81

; TYPE: PRT

; ORGANISM: Agrobacterium tumefaciens

US-09-739-449-10947

Query Match 1.2%; Score 6; DB 5; Length 81;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 70 EKPTSA 75
|||||||
Db 5 EKPTSA 10

RESULT 15

US-09-739-449-11088

; Sequence 11088, Application US/09739449
; GENERAL INFORMATION:

; APPLICANT: Hinkle, Gregory J.

; APPLICANT: Slater, Steven C.

; TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof

; FILE REFERENCE: 38-10(15490)C

; CURRENT APPLICATION NUMBER: US/09/739,449

; CURRENT FILING DATE: 2000-12-19

; PRIOR APPLICATION NUMBER: US 09/514,000

; PRIOR FILING DATE: 2000-02-23

; NUMBER OF SEQ ID NOS: 13351

; SEQ ID NO 11088

; LENGTH: 83

; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-09-739-449-11088

Query Match 1.2%; Score 6; DB 5; Length 83;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 LVGLSL 15
| | | | |
Db 9 LVGLSL 14

RESULT 16
US-09-640-211A-805
; Sequence 805, Application US/09640211A

; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; TITLE OF INVENTION: Modification of Gene Transcription
; FILE REFERENCE: 11000.1021CIU
; CURRENT APPLICATION NUMBER: US/09/640.211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 805
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Eucalyptus grandis
US-09-640-211A-805

Query Match 1.2%; Score 6; DB 5; Length 86;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 204 AWTAAE 209
| | | | |
Db 16 AWTAAE 21

RESULT 17
US-09-739-449-11762
; Sequence 11762, Application US/09739449

; GENERAL INFORMATION:
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15490)C
; CURRENT APPLICATION NUMBER: US/09/739,449
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 09/514,000
; PRIOR FILING DATE: 2000-02-23
; NUMBER OF SEQ ID NOS: 13351
; SEQ ID NO 11762
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-09-739-449-11762

Query Match 1.2%; Score 6; DB 5; Length 100;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 261 KVKGTN 266
| | | | |
Db 54 KVKGTN 59

RESULT 18

PCT-US00-35017A-773
; Sequence 773, Application PC/TUS0035017A
; GENERAL INFORMATION:

; APPLICANT: Hyseq Inc
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 784PCT
; CURRENT APPLICATION NUMBER: PCT/US00/35017A
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US09/552,317
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 1478
; SEQ ID NO 773
; LENGTH: 130
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US00-35017A-773

Query Match 1.2%; Score 6; DB 1; Length 130;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 339 SVFRTL 344
| | | | |
Db 104 SVFRTL 109

RESULT 19

PCT-US01-01339-4492
; Sequence 4492, Application PC/TUS0101339
; GENERAL INFORMATION:

; APPLICANT: Human Genome Sciences, Inc., et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006PCT
; CURRENT APPLICATION NUMBER: PCT/US01/01339
; CURRENT FILING DATE: 2001-03-17
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4492
; LENGTH: 138
; TYPE: PRT
; ORGANISM: Homo sapiens

; FEATURE:
; NAME/KEY: SITE
; LOCATION: (32)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (34)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (104)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (105)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (129)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
PCT-US01-01339-4492

Query Match 1.2%; Score 6; DB 1; Length 138;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 313 GCELSG 318
| | | | |
Db 11 GCELSG 16

RESULT 20

PCT-US01-01329-1749
; Sequence 1749, Application PC/TUS0101329
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc., et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P120PCT
; CURRENT APPLICATION NUMBER: PCT/US01/01329
; CURRENT FILING DATE: 2001-01-17
; NUMBER OF SEQ ID NOS: 3506
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1749
; LENGTH: 138
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (32)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (34)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (104)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (105)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (129)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
PCT-US01-01329-1749

Query Match

Best Local Similarity 1.2%; Score 6; DB 1; Length 138;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 313 GCELSG 318

Db 11 GCELSG 16

RESULT 21

US-09-814-666-274
; Sequence 274, Application US/09814666
; GENERAL INFORMATION:
; APPLICANT: Gearing, David P.
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A
; TITLE OF INVENTION: HUMAN PROSTATE EPITHELIUM LIBRARY
; FILE REFERENCE: 1600.1018-001
; CURRENT APPLICATION NUMBER: US/09/814,666
; CURRENT FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: US 09/307,649
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: 60/084,562
; PRIOR FILING DATE: 1998-05-07
; NUMBER OF SEQ ID NOS: 394
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 274
; LENGTH: 145
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)...(22)
; NAME/KEY: VARIANT
; LOCATION: (1)...(145)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-814-666-274

Query Match 1.2%; Score 6; DB 1; Length 145;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query Match

Best Local Similarity 1.2%; Score 6; DB 5; Length 145;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 154 COVKCS 159

Db 63 COVKCS 68

RESULT 22

PCT-US01-01302-91
; Sequence 91, Application PC/TUS0101302
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc., et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT207PCT
; CURRENT APPLICATION NUMBER: PCT/US01/01302
; CURRENT FILING DATE: 2001-01-17
; NUMBER OF SEQ ID NOS: 192
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 91
; LENGTH: 148
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-01302-91

Query Match

Best Local Similarity 1.2%; Score 6; DB 1; Length 148;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 PCRNGG 132

Db 50 PCRNGG 55

RESULT 23

PCT-US01-11988-924
; Sequence 924, Application PC/TUS0111988
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546PCT
; CURRENT APPLICATION NUMBER: PCT/US01/11988
; CURRENT FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: 60/229, 358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256, 931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199, 384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 924
; LENGTH: 186
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-11988-924

Query Match

Best Local Similarity 1.2%; Score 6; DB 1; Length 186;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 LSLGVP 18

Db 79 LSLGVP 84

RESULT 24

US-09-833-245-924

; Sequence 924, Application US/09833245
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546PCT
; CURRENT APPLICATION NUMBER: US/09/833,245
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229, 358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256, 931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199, 384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 924
; LENGTH: 186
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-833-245-924

Query Match 1.2%; Score 6; DB 5; Length 186;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 LSLGVP 18
Db 79 LSLGVP 84

RESULT 25
PCT-US01-01302-109
; Sequence 109, Application PC/TUS0101302
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc., et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT207PCT
; CURRENT APPLICATION NUMBER: PCT/US01/01302
; CURRENT FILING DATE: 2001-01-17
; NUMBER OF SEQ ID NOS: 192
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 109
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (97)
; OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (187)
; OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (188)
; OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
PCT-US01-01302-109

Query Match 1.2%; Score 6; DB 1; Length 188;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 PCRNGG 132
Db 82 PCRNGG 87

Search completed: May 23, 2001, 06:23:48
Job time: 172 sec

GenCore version 4.5
Copyright (C) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 06:15:21 ; Search time 20.57 seconds
(without alignments)
448.284 Million cell updates/sec

Title: US-09-237-981-10

Perfect score: 480
Sequence: 1 MKHLVAAMLLVGLSLGVPQF.....WSWYGRITRLRSELLGCAEEE 480

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 185757 seqs, 19210857 residues

Word size : 0

Total number of hits satisfying chosen parameters: 185757

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 50 summaries

Database : Issued_Patents_AA:*

- 1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/2/1aa/PCTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/2/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	480	100.0	480	2	US-08-480-229C-10
2	480	100.0	480	2	US-08-659-235C-10
3	194	40.4	221	2	US-08-480-229C-29
4	194	40.4	221	2	US-08-659-235C-29
5	79	16.5	513	2	US-08-480-229C-14
6	79	16.5	513	2	US-08-659-235C-14
7	57	11.9	85	2	US-08-480-229C-1
8	57	11.9	85	2	US-08-659-235C-1
9	41	8.5	43	2	US-08-480-229C-24
10	41	8.5	43	2	US-08-659-235C-24
11	40	8.3	321	2	US-08-480-229C-21
12	40	8.3	321	2	US-08-659-235C-21
13	18	3.8	25	2	US-08-480-229C-22
14	18	3.8	25	2	US-08-659-235C-22
15	18	3.8	42	2	US-08-480-229C-25
16	18	3.8	42	2	US-08-659-235C-25
17	14	2.9	57	2	US-08-480-229C-23
18	14	2.9	57	2	US-08-659-235C-23
19	10	2.1	1010	4	US-08-882-046-7
20	10	2.1	1193	3	US-08-400-159-10
21	10	2.1	1193	3	US-08-611-729A-10
22	10	2.1	1218	2	US-08-400-159-6
23	10	2.1	1218	3	US-08-611-729A-6
24	10	2.1	1218	4	US-08-882-046-2
25	10	2.1	1219	4	US-08-882-046-5
26	10	2.1	2556	1	US-08-185-432-17
27	10	2.1	2556	1	US-08-083-590A-20

28	19	2.1	2556	3	US-08-532-384-20	Sequence 20, Appl
29	9	1.9	85	2	US-08-480-229C-4	Sequence 4, Appl
30	9	1.9	85	2	US-08-659-235C-4	Sequence 4, Appl
31	9	1.9	109	1	US-08-111-939-25	Sequence 25, Appl
32	9	1.9	157	2	US-08-162-402B-13	Sequence 13, Appl
33	9	1.9	159	2	US-08-162-402B-12	Sequence 12, Appl
34	9	1.9	217	1	US-07-607-538C-3	Sequence 3, Appl
35	9	1.9	217	1	US-08-162-402B-3	Sequence 3, Appl
36	9	1.9	218	1	US-07-607-538C-2	Sequence 2, Appl
37	9	1.9	218	1	US-08-162-402B-2	Sequence 2, Appl
38	9	1.9	320	2	US-08-480-229C-20	Sequence 20, Appl
39	9	1.9	320	2	US-08-659-235C-20	Sequence 20, Appl
40	9	1.9	387	2	US-08-162-402B-6	Sequence 6, Appl
41	9	1.9	463	2	US-08-162-402B-9	Sequence 9, Appl
42	9	1.9	465	2	US-08-162-402B-8	Sequence 8, Appl
43	9	1.9	923	3	US-08-936-135-6	Sequence 6, Appl
44	9	1.9	1404	3	US-08-400-159-2	Sequence 2, Appl
45	9	1.9	1404	3	US-08-611-729A-2	Sequence 2, Appl
46	9	1.9	2319	1	US-08-212-133A-8	Sequence 8, Appl
47	9	1.9	2319	1	US-08-474-503-6	Sequence 6, Appl
48	9	1.9	2319	2	US-08-670-707A-6	Sequence 6, Appl
49	9	1.9	2319	4	US-09-037-601-6	Sequence 6, Appl
50	9	1.9	2319	5	PCT-US94-13200-6	Sequence 6, Appl

ALIGNMENTS

RESULT 1
US-08-480-229C-10
; Sequence 10, Application US/08480229C
; Patent No. 5874562
; GENERAL INFORMATION:
; APPLICANT: Quettermous, Thomas
; APPLICANT: Hogan, Bridgid
; APPLICANT: Snodgrass, H. Ralph
; APPLICANT: Zupancic, Thomas J.
; TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
; TITLE OF INVENTION: CELL LOCUS-1
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,229C
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-0026-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 Pennie
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 480 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-480-229C-10

```
Query Match 100.0%; Score 480; DB 2; Length 480;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 480; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKHLVAALLVGLSLGVQFGKGDICNPCEGNGICLSGLADDSFSCCEPGEFAGPNC 60
DB 1 MKHLVAALLVGLSLGVQFGKGDICNPCEGNGICLSGLADDSFSCCEPGEFAGPNC 60

QY 61 SYVEVASDEEKPTSGAGPCIPNCHNGGTCEISEAYRGDTFIGYVCKCPRGFNGIHCQNI 120
DB 61 SYVEVASDEEKPTSGAGPCIPNCHNGGTCEISEAYRGDTFIGYVCKCPRGFNGIHCQNI 120

QY 121 NECEAPCRNGGICTDLVANYSCCEPGEFGRNCQYKCSGHLGIEGGIISNOQITASSNH 180
DB 121 NECEAPCRNGGICTDLVANYSCCEPGEFGRNCQYKCSGHLGIEGGIISNOQITASSNH 180

QY 181 RALFGLQKWYPYARLNKKGLINAWTAANDRWPFQIQLNQRKMRVTGVTGAKRIGSP 240
DB 181 RALFGLQKWYPYARLNKKGLINAWTAANDRWPFQIQLNQRKMRVTGVTGAKRIGSP 240

QY 241 EYIKSYKTAISNDGKTWAMYKVGKTNEEMVFRGNVDNNTPYANSTFTPPPIKAQYVRLYPOI 300
DB 241 EYIKSYKTAISNDGKTWAMYKVGKTNEEMVFRGNVDNNTPYANSTFTPPPIKAQYVRLYPOI 300

QY 301 CRRHCTLRMELLCGELSCSEPLGMKSGHIGDYQITASSVFTLNMDMFTWEPKARLDK 360
DB 301 CRRHCTLRMELLCGELSCSEPLGMKSGHIGDYQITASSVFTLNMDMFTWEPKARLDK 360

QY 361 QGKVNWATSGHNDOSQWLQVLLVPTKVTGIITQGAADFGHGVQFVGSYKLAISNDGEHWM 420
DB 361 QGKVNWATSGHNDOSQWLQVLLVPTKVTGIITQGAADFGHGVQFVGSYKLAISNDGEHWM 420

QY 421 VHODEKQKQKVKVQGNFNDNTHRNKNVDPPPIYARFIRILPWSWYGRITLRSSELLGCASEE 480
DB 421 VHODEKQKQKVKVQGNFNDNTHRNKNVDPPPIYARFIRILPWSWYGRITLRSSELLGCASEE 480

RESULT 2
US-08-659-235C-10
; Sequence 10, Application US/08659235C
; Patent No. 5877281
; GENERAL INFORMATION:
; APPLICANT: Quertermous, Thomas
; APPLICANT: Hogan, Brigid
; APPLICANT: Snodgrass, H. Ralph
; APPLICANT: Zupancic, Thomas J.
; TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/659,235C
; FILING DATE: 05-JUN-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-0034-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 Pennie
```

```
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
; LENGTH: 480 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-659-235C-10

Query Match 100.0%; Score 480; DB 2; Length 480;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 480; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKHLVAALLVGLSLGVQFGKGDICNPCEGNGICLSGLADDSFSCCEPGEFAGPNC 60
DB 1 MKHLVAALLVGLSLGVQFGKGDICNPCEGNGICLSGLADDSFSCCEPGEFAGPNC 60

QY 61 SYVEVASDEEKPTSGAGPCIPNCHNGGTCEISEAYRGDTFIGYVCKCPRGFNGIHCQNI 120
DB 61 SYVEVASDEEKPTSGAGPCIPNCHNGGTCEISEAYRGDTFIGYVCKCPRGFNGIHCQNI 120

QY 121 NECEAPCRNGGICTDLVANYSCCEPGEFGRNCQYKCSGHLGIEGGIISNOQITASSNH 180
DB 121 NECEAPCRNGGICTDLVANYSCCEPGEFGRNCQYKCSGHLGIEGGIISNOQITASSNH 180

QY 181 RALFGLQKWYPYARLNKKGLINAWTAANDRWPFQIQLNQRKMRVTGVTGAKRIGSP 240
DB 181 RALFGLQKWYPYARLNKKGLINAWTAANDRWPFQIQLNQRKMRVTGVTGAKRIGSP 240

QY 241 EYIKSYKTAISNDGKTWAMYKVGKTNEEMVFRGNVDNNTPYANSTFTPPPIKAQYVRLYPOI 300
DB 241 EYIKSYKTAISNDGKTWAMYKVGKTNEEMVFRGNVDNNTPYANSTFTPPPIKAQYVRLYPOI 300

QY 301 CRRHCTLRMELLCGELSCSEPLGMKSGHIGDYQITASSVFTLNMDMFTWEPKARLDK 360
DB 301 CRRHCTLRMELLCGELSCSEPLGMKSGHIGDYQITASSVFTLNMDMFTWEPKARLDK 360

QY 361 QGKVNWATSGHNDOSQWLQVLLVPTKVTGIITQGAADFGHGVQFVGSYKLAISNDGEHWM 420
DB 361 QGKVNWATSGHNDOSQWLQVLLVPTKVTGIITQGAADFGHGVQFVGSYKLAISNDGEHWM 420

QY 421 VHODEKQKQKVKVQGNFNDNTHRNKNVDPPPIYARFIRILPWSWYGRITLRSSELLGCASEE 480
DB 421 VHODEKQKQKVKVQGNFNDNTHRNKNVDPPPIYARFIRILPWSWYGRITLRSSELLGCASEE 480

RESULT 3
US-08-480-229C-29
; Sequence 29, Application US/08480229C
; Patent No. 5874562
; GENERAL INFORMATION:
; APPLICANT: Quertermous, Thomas
; APPLICANT: Hogan, Brigid
; APPLICANT: Snodgrass, H. Ralph
; APPLICANT: Zupancic, Thomas J.
; TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,229C
; FILING DATE: 07-JUN-1995
```

```
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-0026-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 221 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-480-229C-29

Query Match 40.4%; Score 194; DB 2; Length 221;
Best Local Similarity 100.0%; Pred. No. 2.5e-180;
Matches 194; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKHLVAALLVGLSLGVPGFGKGDICNPNCENGICLSGLADDSFSCCEPGFAGPNC 60
Db 1 MKHLVAALLVGLSLGVPGFGKGDICNPNCENGICLSGLADDSFSCCEPGFAGPNC 60
QY 61 SVVEVASDEEKTPTAGPCIPNCHNGGTCETSEAYRGDTFTGYVCKCPRGFNGIHCOHNI 120
Db 61 SVVEVASDEEKTPTAGPCIPNCHNGGTCETSEAYRGDTFTGYVCKCPRGFNGIHCOHNI 120
QY 121 NECEAEPCRNCGICTDLVANYSCCEPGFPMGRNCOYKCSGHLGIEGGIISNOQITASSNH 180
Db 121 NECEAEPCRNCGICTDLVANYSCCEPGFPMGRNCOYKCSGHLGIEGGIISNOQITASSNH 180
QY 181 RALFGLQKWYPYA 194
Db 181 RALFGLQKWYPYA 194

RESULT 4
US-08-659-235C-29
Sequence 29, Application US/08659235C
Patent No. 5877281
GENERAL INFORMATION:
APPLICANT: Quertemous, Thomas
APPLICANT: Hogan, Brigid
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
TITLE OF INVENTION: CELL LOCUS-1
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/659,235C
FILING DATE: 05-JUN-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-0034-999
```

```
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennie
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 221 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-659-235C-29

Query Match 40.4%; Score 194; DB 2; Length 221;
Best Local Similarity 100.0%; Pred. No. 2.5e-180;
Matches 194; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKHLVAALLVGLSLGVPGFGKGDICNPNCENGICLSGLADDSFSCCEPGFAGPNC 60
Db 1 MKHLVAALLVGLSLGVPGFGKGDICNPNCENGICLSGLADDSFSCCEPGFAGPNC 60
QY 61 SVVEVASDEEKTPTAGPCIPNCHNGGTCETSEAYRGDTFTGYVCKCPRGFNGIHCOHNI 120
Db 61 SVVEVASDEEKTPTAGPCIPNCHNGGTCETSEAYRGDTFTGYVCKCPRGFNGIHCOHNI 120
QY 121 NECEAEPCRNCGICTDLVANYSCCEPGFPMGRNCOYKCSGHLGIEGGIISNOQITASSNH 180
Db 121 NECEAEPCRNCGICTDLVANYSCCEPGFPMGRNCOYKCSGHLGIEGGIISNOQITASSNH 180
QY 181 RALFGLQKWYPYA 194
Db 181 RALFGLQKWYPYA 194

RESULT 5
US-08-480-229C-14
Sequence 14, Application US/08480229C
Patent No. 5874562
GENERAL INFORMATION:
APPLICANT: Quertemous, Thomas
APPLICANT: Hogan, Brigid
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
TITLE OF INVENTION: CELL LOCUS-1
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,229C
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-0026-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennie
INFORMATION FOR SEQ ID NO: 14:
```

;; SEQUENCE CHARACTERISTICS:
; LENGTH: 513 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-480-229C-14

Query Match 16.5%; Score 79; DB 2; Length 513;
Best Local Similarity 100.0%; Pred. No. 1.5e-68;
Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 341 FRTLNDMFTWEPKRLDKQKGVNAWTSQHNDQSQWLQVDLLVPTKVTGIIITQAKDFG 400
|||||
Db 374 FRTLNDMFTWEPKRLDKQKGVNAWTSQHNDQSQWLQVDLLVPTKVTGIIITQAKDFG 433

Qy 401 HVQFVGSYKLAYSNDGEHW 419
|||||
Db 434 HVQFVGSYKLAYSNDGEHW 452

RESULT 6

US-08-659-235C-14
; Sequence 14, Application US/08659235C
; Patent No. 5877281

;; GENERAL INFORMATION:
; APPLICANT: Quertermous, Thomas
; APPLICANT: Hogan, Brigid
; APPLICANT: Snodgrass, H. Ralph
; APPLICANT: Zupancic, Thomas J.
; TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2711

;; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/659,235C
; FILING DATE: 05-JUN-1996

;; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-0034-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 Pennie

;; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 513 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-659-235C-14

Query Match 16.5%; Score 79; DB 2; Length 513;
Best Local Similarity 100.0%; Pred. No. 1.5e-68;
Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 341 FRTLNDMFTWEPKRLDKQKGVNAWTSQHNDQSQWLQVDLLVPTKVTGIIITQAKDFG 400

Db 374 FRTLNDMFTWEPKRLDKQKGVNAWTSQHNDQSQWLQVDLLVPTKVTGIIITQAKDFG 433
Qy 401 HVQFVGSYKLAYSNDGEHW 419
|||||
Db 434 HVQFVGSYKLAYSNDGEHW 452

RESULT 7

US-08-480-229C-1
; Sequence 1, Application US/08480229C
; Patent No. 5874562

;; GENERAL INFORMATION:
; APPLICANT: Quertermous, Thomas
; APPLICANT: Hogan, Brigid
; APPLICANT: Snodgrass, H. Ralph
; APPLICANT: Zupancic, Thomas J.
; TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2711

;; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,229C
; FILING DATE: 07-JUN-1995

;; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-0026-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 Pennie
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 85 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-480-229C-1

Query Match 11.9%; Score 57; DB 2; Length 85;
Best Local Similarity 100.0%; Pred. No. 6.3e-48;
Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 404 FVGSYKLAYSNDGEHWMVHODEKQKDKVFGNFDNTHRKNVDPPIYARILP 460
|||||
Db 26 FVGSYKLAYSNDGEHWMVHODEKQKDKVFGNFDNTHRKNVDPPIYARILP 82

RESULT 8

US-08-659-235C-1
; Sequence 1, Application US/08659235C
; Patent No. 5877281

;; GENERAL INFORMATION:
; APPLICANT: Quertermous, Thomas
; APPLICANT: Hogan, Brigid
; APPLICANT: Snodgrass, H. Ralph
; APPLICANT: Zupancic, Thomas J.
; TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL

;; TITLE OF INVENTION: CELL LOCUS-1
;; NUMBER OF SEQUENCES: 29
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Pennie & Edmonds LLP
;; STREET: 1155 Avenue of the Americas
;; CITY: New York
;; STATE: New York
;; COUNTRY: United States
;; ZIP: 10036-2711
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/659,235C
;; FILING DATE: 05-JUN-1996
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Poissant, Brian M.
;; REGISTRATION NUMBER: 28,462
;; REFERENCE/DOCKET NUMBER: 8907-0034-999
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (212) 790-9090
;; TELEX: 66141 Pennie
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 85 amino acids
;; TYPE: amino acid
;; STRANDEDNESS:
;; TOPOLOGY: unknown
;; MOLECULE TYPE: peptide
;; US-08-659-235C-1

Query Match 11.9%; Score 57; DB 2; Length 85;
Best Local Similarity 100.0%; Pred. No. 6.3e-48;
Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 404 FVGSYKLAYSNDGEHMHVHDEKQKDKVFGNFDNDTHRNKVIDPPIYARFILP 460
DB 26 FVGSYKLAYSNDGEHMHVHDEKQKDKVFGNFDNDTHRNKVIDPPIYARFILP 82

RESULT 9
US-08-480-229C-24
; Sequence 24, Application US/08480229C
; Patent No. 5874562
; GENERAL INFORMATION:
; APPLICANT: Quettermous, Thomas
; APPLICANT: Hogan, Brigid
; APPLICANT: Snodgrass, H. Ralph
; APPLICANT: Zupancic, Thomas J.
; TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2711
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,229C
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 536

;; ATTORNEY/AGENT INFORMATION:
;; NAME: Poissant, Brian M.
;; REGISTRATION NUMBER: 28,462
;; REFERENCE/DOCKET NUMBER: 8907-0026-999
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (212) 790-9090
;; TELEX: 66141 Pennie
;; INFORMATION FOR SEQ ID NO: 24:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 43 amino acids
;; TYPE: amino acid
;; STRANDEDNESS:
;; TOPOLOGY: unknown
;; MOLECULE TYPE: peptide
;; US-08-480-229C-24

Query Match 8.5%; Score 41; DB 2; Length 43;
Best Local Similarity 100.0%; Pred. No. 1.1e-32;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 PNPCHNGTGTCEISEAYRGDTFFIGYVCKCPRGFNGIHCOHNI 120
DB 3 PNPCHNGTGTCEISEAYRGDTFFIGYVCKCPRGFNGIHCOHNI 43

RESULT 10
US-08-659-235C-24
; Sequence 24, Application US/08659235C
; Patent No. 5877281
; GENERAL INFORMATION:
; APPLICANT: Quettermous, Thomas
; APPLICANT: Hogan, Brigid
; APPLICANT: Snodgrass, H. Ralph
; APPLICANT: Zupancic, Thomas J.
; TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2711
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/659,235C
; FILING DATE: 05-JUN-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-0034-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEX: 66141 Pennie
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 43 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; US-08-659-235C-24

APPLICANT: SHODGRASS, H. Ralph
APPLICANT: ZUPARSKI, Thomas J.

APPLICATION NUMBER: US/

US-08-480-229C-21

APPLICANT: SHODGRASS, H. Ralph
APPLICANT: ZUPARSKI, Thomas J.

CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-0026-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennie
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 25 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-480-229C-22

Query Match 3.8%; Score 18; DB 2; Length 25;
Best Local Similarity 100.0%; Pred. No. 1.3e-10;
Matches 18; Conservative 0; Mismatches 0; Indels 0;

QY 8 WLLVGLSLGVPQFGKGDI 25
|||||
Db 8 WLLVGLSLGVPQFGKGDI 25

RESULT 14
US-08-659-235C-22
Sequence 22, Application US/08659235C
Patent No. 5877281
GENERAL INFORMATION:
APPLICANT: Quettermous, Thomas
APPLICANT: Hogan, Brigid
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/659,235C
FILING DATE: 05-JUN-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-0034-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennie
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 25 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-659-235C-22

Query Match 3.8%; Score 18; DB 2; Length 25;
Best Local Similarity 100.0%; Pred. No. 1.3e-10;
Matches 18; Conservative 0; Mismatches 0; Indels 0;

QY 8 WLLVGLSLGVPQFGKGDI 25
|||||
Db 8 WLLVGLSLGVPQFGKGDI 25

RESULT 15
US-08-480-229C-25
Sequence 25, Application US/08480229C
Patent No. 5874562
GENERAL INFORMATION:
APPLICANT: Quettermous, Thomas
APPLICANT: Hogan, Brigid
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,229C
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-0026-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennie
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 42 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-480-229C-25

Query Match 3.8%; Score 18; DB 2; Length 42;
Best Local Similarity 100.0%; Pred. No. 2.1e-10;
Matches 18; Conservative 0; Mismatches 0; Indels 0;

QY 140 NYSCECPGEFMGRNCQYK 157
|||||
Db 25 NYSCECPGEFMGRNCQYK 42

RESULT 16
US-08-659-235C-25
Sequence 25, Application US/08659235C
Patent No. 5877281
GENERAL INFORMATION:
APPLICANT: Quettermous, Thomas
APPLICANT: Hogan, Brigid
APPLICANT: Snodgrass, H. Ralph

APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
CELL LOCUS-1
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/659,235C
FILING DATE: 05-JUN-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-0034-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennie
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 42 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-659-235C-25

Query Match 3.8%; Score 18; DB 2; Length 42;
Best Local Similarity 100.0%; Pred. No. 2.1e-10;
Matches 18; Conservative 0; Mismatches 0; Indels 0;

QY 140 NYSCECPGFMRNCQYK 157
Db 25 NYSCECPGFMRNCQYK 42

RESULT 17
US-08-480-229C-23
Sequence 23, Application US/08480229C
Patent No. 5874562
GENERAL INFORMATION:
APPLICANT: Quettermous, Thomas
APPLICANT: Hogan, Bridgid
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
CELL LOCUS-1
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,229C

FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-0026-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennie
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 57 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-480-229C-23

Query Match 2.9%; Score 14; DB 2; Length 57;
Best Local Similarity 100.0%; Pred. No. 2.1e-06;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 PNCSSVVEVASDEE 70
Db 37 PNCSSVVEVASDEE 50

RESULT 18
US-08-659-235C-23
Sequence 23, Application US/08659235C
Patent No. 5877281
GENERAL INFORMATION:
APPLICANT: Quettermous, Thomas
APPLICANT: Hogan, Bridgid
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
CELL LOCUS-1
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/659,235C
FILING DATE: 05-JUN-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-0034-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennie
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 57 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-659-235C-23

Query Match 2.9%; Score 14; DB 2; Length 57;
Best Local Similarity 100.0%; Pred. No. 2.1e-06;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 PNCSSVVEVASDEE 70
|||||

Db 37 PNCSSVVEVASDEE 50

RESULT 19

US-08-882-046-7
; Sequence 7, Application US/08882046
; Patent No. 6136952

GENERAL INFORMATION:

; APPLICANT: Li, Linheng
; APPLICANT: Hood, Leroy
; APPLICANT: Krantz, Ian D.
; APPLICANT: Spinner, Nancy B.
; TITLE OF INVENTION: Human Jagged Polypeptide, Encoding
; Nucleic Acids and Methods of Use
; NUMBER OF SEQUENCES: 110
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122

COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/882.046
; FILING DATE: 25-JUN-1997
; CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:

; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-UW 2637
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1010 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein

US-08-882-046-7

Query Match 2.1%; Score 10; DB 4; Length 1010;
Best Local Similarity 100.0%; Pred. No. 0.2;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 PNPCHNGGTC 89
|||||

Db 750 PNPCHNGGTC 759

RESULT 20

US-08-400-159-10
; Sequence 10, Application US/08400159
; Patent No. 5869282

GENERAL INFORMATION:

; APPLICANT: Ish-Horowicz, David
; APPLICANT: Henrique, Domingos M.P.
; APPLICANT: Lewis, Julian H.
; APPLICANT: Myat, Anna M.

; APPLICANT: Fleming, Robert J.
; APPLICANT: Artavanis-Tsakonas, Spyridon
; APPLICANT: Mann, Robert S.
; APPLICANT: Gray, Grace E.
; TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF THE
; SERRATE GENE AND METHODS BASED THEREON
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711

COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/400,159
; FILING DATE: 07-MAR-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mirock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7326-029
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1193 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-400-159-10

Query Match 2.1%; Score 10; DB 2; Length 1193;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 PNPCHNGGTC 89
|||||

Db 724 PNPCHNGGTC 733

RESULT 21

US-08-611-729A-10
; Sequence 10, Application US/08611729A
; Patent No. 6004924

GENERAL INFORMATION:

; APPLICANT: Ish-Horowicz, David
; APPLICANT: Henrique, Domingos M.P.
; APPLICANT: Lewis, Julian H.
; APPLICANT: Myat, Anna M.
; APPLICANT: Fleming, Robert J.
; APPLICANT: Artavanis-Tsakonas, Spyridon
; APPLICANT: Mann, Robert S.
; APPLICANT: Gray, Grace E.

; TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF THE
; SERRATE GENE AND METHODS BASED THEREON
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711

COMPUTER READABLE FORM:

```
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/611,729A
; FILING DATE: 06-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7326-037
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1193 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-611-729A-10

Query Match          2.1%; Score 10; DB 3; Length 1193;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 PNPCHNGGTC 89
Db 724 PNPCHNGGTC 733

RESULT 22
US-08-400-159-6
; Sequence 6, Application US/08400159
; Patent No. 5869282
; GENERAL INFORMATION:
; APPLICANT: Ish-Horowicz, David
; APPLICANT: Henrique, Domingos M.P.
; APPLICANT: Lewis, Julian H.
; APPLICANT: Myat, Anna M.
; APPLICANT: Fleming, Robert J.
; APPLICANT: Artavanis-Tsakonas, Spyridon
; APPLICANT: Mann, Robert S.
; APPLICANT: Gray, Grace E.
; TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF THE
; TITLE OF INVENTION: SERRATE GENE AND METHODS BASED THEREON
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/400,159
; FILING DATE: 07-MAR-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7326-029
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
```

```
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1218 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-400-159-6

Query Match          2.1%; Score 10; DB 2; Length 1218;
Best Local Similarity 100.0%; Pred. No. 0.24;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 PNPCHNGGTC 89
Db 750 PNPCHNGGTC 759

RESULT 23
US-08-611-729A-6
; Sequence 6, Application US/08611729A
; Patent No. 6004924
; GENERAL INFORMATION:
; APPLICANT: Ish-Horowicz, David
; APPLICANT: Henrique, Domingos M.P.
; APPLICANT: Lewis, Julian H.
; APPLICANT: Myat, Anna M.
; APPLICANT: Fleming, Robert J.
; APPLICANT: Artavanis-Tsakonas, Spyridon
; APPLICANT: Mann, Robert S.
; APPLICANT: Gray, Grace E.
; TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF THE
; TITLE OF INVENTION: SERRATE GENE AND METHODS BASED THEREON
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/611,729A
; FILING DATE: 06-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7326-037
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1218 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-611-729A-6

Query Match          2.1%; Score 10; DB 3; Length 1218;
Best Local Similarity 100.0%; Pred. No. 0.24;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 PNPCHNGGTC 89
```

Db 750 PNPCHGGTC 759
|||||

RESULT 24

US-08-882-046-2
; Sequence 2, Application US/08882046
; Patent No. 6136952
; GENERAL INFORMATION:
; APPLICANT: Li, Linheng
; APPLICANT: Hood, Leroy
; APPLICANT: Krantz, Ian D.
; APPLICANT: Spinner, Nancy B.
; TITLE OF INVENTION: Human Jagged Polypeptide, Encoding
; TITLE OF INVENTION: Nucleic Acids and Methods of Use
; NUMBER OF SEQUENCES: 110
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/882,046
; FILING DATE: 25-JUN-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-UW 2637
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1218 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-882-046-2

Query Match 2.1%; Score 10; DB 4; Length 1218;
Best Local Similarity 100.0%; Pred. No. 0.24;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 PNPCHGGTC 89
|||||

Db 750 PNPCHGGTC 759

RESULT 25

US-08-882-046-5
; Sequence 5, Application US/08882046
; Patent No. 6136952
; GENERAL INFORMATION:
; APPLICANT: Li, Linheng
; APPLICANT: Hood, Leroy
; APPLICANT: Krantz, Ian D.
; APPLICANT: Spinner, Nancy B.
; TITLE OF INVENTION: Human Jagged Polypeptide, Encoding
; TITLE OF INVENTION: Nucleic Acids and Methods of Use
; NUMBER OF SEQUENCES: 110
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego

; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/882,046
; FILING DATE: 25-JUN-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-UW 2637
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1219 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-882-046-5

Query Match 2.1%; Score 10; DB 4; Length 1219;
Best Local Similarity 100.0%; Pred. No. 0.24;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 PNPCHGGTC 89
|||||

Db 750 PNPCHGGTC 759

Search completed: May 23, 2001, 06:21:42
Job time: 381 sec

GenCore version 4.5
Copyright (C) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 06:22:17 ; Search time 28.74 Seconds
(without alignments)

528,454 Million cell updates/sec

Title: US-09-237-981-29

Perfect score: 221

Sequence: 1 MKHLVAALLVGLSLGVPQF.....INAWTAENDRNPWIVQTVTG 221

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 198801 seqs, 68722935 residues

Word size : 0

Total number of hits satisfying chosen parameters: 198801

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 50 summaries

Database :

PIR.67.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	10	4.5	1220	2 A56136	jagged protein pre
2	10	4.5	2555	2 A40043	notch protein homo
3	9	4.1	1408	2 S16148	gene serrate prote
4	9	4.1	2139	2 A35672	crumbs protein - f
5	9	4.1	2437	2 S42612	transmembrane prot
6	9	4.1	2531	2 S18188	notch protein homo
7	9	4.1	2531	2 A46019	Notch-1 protein -
8	8	3.6	63	2 F75371	hypothetical prote
9	8	3.6	247	2 E75609	amino acid ABC tra
10	8	3.6	330	2 T46256	brevican - human (
11	8	3.6	481	2 T48516	probable oligopept
12	8	3.6	832	2 A31246	neurogenic protein
13	8	3.6	833	2 S19087	gene Delta protein
14	8	3.6	880	2 S00670	neurogenic repetit
15	8	3.6	883	2 S49126	brevican precursor
16	8	3.6	883	2 S57653	brevican precursor
17	8	3.6	912	2 A54423	brevican precursor
18	8	3.6	1193	2 T21133	hypothetical prote
19	8	3.6	1251	2 A57293	latent transformin
20	8	3.6	1438	2 A48216	neurexin III-alpha
21	8	3.6	1471	2 B48218	neurexin III-alpha
22	8	3.6	1578	2 I48216	neurexin II-alpha
23	8	3.6	1715	2 C40228	notch homolog - Af
24	8	3.6	2524	2 A35844	notch homolog - se
25	8	3.6	2531	2 T31070	GCN1 homolog - fis
26	8	3.6	2670	2 T37919	hypothetical prote
27	7	3.2	106	2 A29760	50S ribosomal prot
28	7	3.2	107	2 C81232	hypothetical prote
29	7	3.2	126	2 B83265	hypothetical prote

30	7	3.2	135	2 C70890	hypothetical prote
31	7	3.2	149	1 CYFGA2	alpha-crystallin c
32	7	3.2	167	1 CYFGAA	hypothetical prote
33	7	3.2	172	2 B71464	conserved hypotet
34	7	3.2	172	2 C81726	alpha-crystallin c
35	7	3.2	173	1 CYHAAA	alpha-crystallin c
36	7	3.2	173	1 CYCHAA	alpha-crystallin c
37	7	3.2	173	1 CYEHAA	alpha-crystallin c
38	7	3.2	173	1 CYAQAA	alpha-crystallin c
39	7	3.2	173	1 CYLZAA	alpha-crystallin c
40	7	3.2	173	2 JC4148	alpha-crystallin -
41	7	3.2	259	2 S53400	RING finger protei
42	7	3.2	280	2 F75057	hypothetical prote
43	7	3.2	291	2 I38098	t-plasminogen acti
44	7	3.2	308	2 S40873	1,4-dihydroxy-2-na
45	7	3.2	320	2 T18732	hypothetical prote
46	7	3.2	328	2 T36494	probable membrane
47	7	3.2	349	2 A34815	carcinoembryonic a
48	7	3.2	382	2 T04260	hypothetical prote
49	7	3.2	402	2 A46089	UDPglucose 6-dehyd
50	7	3.2	437	2 D83329	conserved hypotet

ALIGNMENTS

RESULT 1

A56136
jagged protein precursor - rat

C:Species: Rattus norvegicus (Norway rat)

C>Date: 28-Apr-1995 #sequence_revision 28-Apr-1995 #text_change 11-Jan-2000

C:Accession: A56136

R:Rindseell, C.E.; Shawber, C.J.; Boulter, J.; Weinmaster, G.

Cell 80, 909-917, 1995

A:Title: Jagged, a mammalian ligand that activates Notch1.

A:Reference number: A56136; MUID:95211842

A:Accession: A56136

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1220 <LIN>

A:Cross-references: GB:L38483

C:Superfamily: unassigned EGF-related proteins; EGF homology

F:379-410/Domain: EGF homology <EGF1>

F:492-523/Domain: EGF homology <EGF>

F:634-665/Domain: EGF homology <EGF2>

Query Match 4.5% Score 10: DB 2: Length 1220;
Best Local Similarity 100.0%; Pred. No. 0.034; 0; Mismatches 0; Indels 0; Gaps 0;
Matches 10; Conservative 0;

QY 80 PNPCHNGGTC 89

|||||

Db 751 PNPCHNGGTC 760

RESULT 2

A40043

notch protein homolog TAN-1 precursor - human

C:Species: Homo sapiens (man)

C>Date: 21-Apr-1992 #sequence_revision 21-Apr-1992 #text_change 13-Aug-1999

C:Accession: A40043

R:Ellisen, L.W.; Bird, J.; West, D.C.; Soreng, A.L.; Reynolds, T.C.; Smith, S.D.; Skl

Cell 66, 649-661, 1991

A:Title: TAN-1, the human homolog of the Drosophila Notch gene, is broken by chromoso

A:Reference number: A40043; MUID:91347367

A:Accession: A40043

A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual

A:Molecule type: mRNA

A:Residues: 1-2555 <ELL>

A:Cross-references: GB:M73980

C:Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homol

F:261-292/Domain: EGF homology <EGX1>

```

F:494-525/Domain: EGF homology <EGF1>
F:987-1018/Domain: EGF homology <EGX2>
F:1149-1180/Domain: EGF homology <EGF>
F:1187-1218/Domain: EGF homology <EGF3>
F:1233-1264/Domain: EGF homology <EGF3>
F:1927-1959/Domain: ankyrin repeat homology <AN1>
F:1960-1992/Domain: ankyrin repeat homology <AN2>
F:1994-2026/Domain: ankyrin repeat homology <AN3>
F:2027-2059/Domain: ankyrin repeat homology <AN4>
F:2060-2092/Domain: ankyrin repeat homology <AN5>

Query Match          4.58; Score 10; DB 2; Length 2555;
Best Local Similarity 100.08; Pred. No. 0.065;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 NPCHNGTCE 90
Db 684 NPCHNGTCE 693
|||||

RESULT 3
SI6148
gene serrate protein precursor - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C>Date: 31-Dec-1991 #sequence_revision 02-Aug-1994 #text_change 17-Nov-2000
C:Accession: SI6148; SI6878; A36666
R:Thomas, U.; Speicher, S. A.; Knust, E.
Development 111, 749-761, 1991
A:Title: The Drosophila gene Serrate encodes an EGF-like transmembrane protein with a co
A:Reference number: SI6148; MUID:91347903
A:Accession: SI6148
A:Molecule type: mRNA
A:Residues: 1-1408 <THO1>
A:Cross-references: EMBL:X56811
R:Thomas, U.
submitted to the EMBL Data Library, November 1990
A:Reference number: SI6878
A:Accession: SI6878
A:Molecule type: mRNA
A:Residues: 1-1351, 'T', 1353-1408 <THO2>
A:Cross-references: EMBL:X56811; NID:g8563; PID:g8564
R:Fleming, R. J.; Scottgale, T. N.; Diederich, R. J.; Artavanis-Tsakonas, S.
Genes Dev. 4, 2188-2201, 1990
A:Title: The gene Serrate encodes a putative EGF-like transmembrane protein essential fo
A:Reference number: A36666; MUID:91099666
A:Accession: A36666
A>Status: Preliminary
A:Molecule type: mRNA
A:Residues: 1-15, 20-26, 'A', 28-1408 <FLE>
A:Cross-references: GB:M35759; NID:g158605; PID:g158606
C:Genetics:
A:Gene: FlyBase:Ser
A:Cross-references: FlyBase:FBgn0004197
C:Superfamily: unassigned EGF-related proteins; EGF homology
C:Keywords: glycoprotein; transmembrane protein
F:1-84/Domain: signal sequence #status predicted <SIG>
F:85-1408/Product: gene serrate protein #status predicted <MAT>
F:85-1221/Domain: extracellular #status predicted <EXT>
F:283-316/Domain: EGF homology <EG01>
F:319-348/Domain: EGF homology <EG02>
F:355-388/Domain: EGF homology <EG03>
F:395-488/Domain: EGF homology #status atypical <EG04>
F:495-526/Domain: EGF homology <EG05>
F:533-608/Domain: EGF homology #status atypical <EG06>
F:615-645/Domain: EGF homology <EG07>
F:652-683/Domain: EGF homology <EG08>
F:690-720/Domain: EGF homology <EG09>
F:727-796/Domain: EGF homology #status atypical <EG10>
F:803-834/Domain: EGF homology <EG11>
F:841-876/Domain: EGF homology <EG12>
F:883-914/Domain: EGF homology <EG13>
F:921-952/Domain: EGF homology <EG14>

F:997-1060/Region: cysteine-rich
F:1222-1246/Domain: transmembrane #status predicted <TM1>
F:1247-1408/Domain: intracellular #status predicted <INT>
F:152-196,247,331,412,452,558,739,965,977,1004,1030,1150/Binding site: carbohydrate (

Query Match          4.18; Score 9; DB 2; Length 1408;
Best Local Similarity 100.08; Pred. No. 0.47;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 82 PCHNGTCE 90
Db 887 PCHNGTCE 895
|||||

RESULT 4
A35672
crumbs protein - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C>Date: 21-Sep-1990 #sequence_revision 18-Nov-1992 #text_change 11-Jan-2000
C:Accession: A35672
R:Repas, U.; Therès, C.; Knust, E.
Cell 61, 787-799, 1990
A:Title: crumbs encodes an EGF-like protein expressed on apical membranes of Drosophi
A:Reference number: A35672; MUID:90263104
A:Accession: A35672
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-2139 <TEP>
A:Cross-references: GB:M33753
A:Note: the authors translated the codon GGC for residue 1928 as Cys, and TAT for res
C:Genetics:
A:Gene: FlyBase:crb
A:Cross-references: FlyBase:FBgn0000368
C:Superfamily: unassigned EGF-related proteins; EGF homology
C:Keywords: transmembrane protein
F:352-385/Domain: EGF homology <EGX1>
F:392-424/Domain: EGF homology <EGF1>
F:691-722/Domain: EGF homology <EGF>
F:767-799/Domain: EGF homology <EGF3>
F:1878-1914/Domain: EGF homology <EGX2>

Query Match          4.18; Score 9; DB 2; Length 2139;
Best Local Similarity 100.08; Pred. No. 0.67;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 48 CECPEGFAG 56
Db 2058 CECPEGFAG 2066
|||||

RESULT 5
S42612
transmembrane protein precursor - zebra fish
C:Species: Brachydanio rerio (zebra fish)
C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 20-Sep-1999
C:Accession: S42612
R:Bierkamp, C.; Campos-Ortega, J. A.
Mech. Dev. 43, 87-100, 1993
A:Title: A zebrafish homologue of the Drosophila neurogenic gene Notch and its patter
A:Reference number: S42612; MUID:94128602
A:Accession: S42612
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-2437 <BIE>
A:Cross-references: EMBL:X69088; NID:g433866; PIDN:CAA48831.1; PTD:g433867
C:Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homol
F:755-786/Domain: EGF homology <EGF1>
F:1023-1054/Domain: EGF homology <EGF2>
F:1185-1216/Domain: EGF homology <EGF2>
F:1915-1947/Domain: ankyrin repeat homology <AN1>
F:1948-1980/Domain: ankyrin repeat homology <AN2>

```

F:1982-2014/Domain: ankyrin repeat homology <AN3>
F:2015-2047/Domain: ankyrin repeat homology <AN4>
F:2048-2080/Domain: ankyrin repeat homology <AN5>

Query Match 4.1%; Score 9; DB 2; Length 2437;
Best Local Similarity 100.0%; Pred. No. 0.75;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 NPCHNGTGC 89
|||||
DB 683 NPCHNGTGC 691

RESULT 6

notch protein homolog - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 19-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 13-Aug-1999
C:Accession: S18188
R:Weinmaster, G.; Roberts, V.J.; Lemke, G.
Development 113, 199-205, 1991
A:Title: A homolog of Drosophila Notch expressed during mammalian development.
A:Reference number: S18188; MUID:92111383
A:Accession: S18188
A:Molecule type: mRNA
A:Residues: 1-2531 <WEI>
A:Cross-references: EMBL:X57405; NID:g57634; PID:g57635
C:Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homology
F:987-1018/Domain: EGF homology <EGF1>
F:1025-1056/Domain: EGF homology <EGF2>
F:1233-1264/Domain: EGF homology <EGF3>
F:1917-1949/Domain: ankyrin repeat homology <AN1>
F:1950-1982/Domain: ankyrin repeat homology <AN2>
F:1984-2016/Domain: ankyrin repeat homology <AN3>
F:2017-2049/Domain: ankyrin repeat homology <AN4>
F:2050-2082/Domain: ankyrin repeat homology <AN5>

Query Match 4.1%; Score 9; DB 2; Length 2531;
Best Local Similarity 100.0%; Pred. No. 0.77;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 82 PCHNGTGC 90
|||||
DB 686 PCHNGTGC 694

RESULT 7

Notch-1 protein - mouse
N:Alternate names: notch protein
C:Species: Mus musculus (house mouse)
C:Date: 22-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 20-Sep-1999
C:Accession: A46019; S25144
R:del Amo, F.F.; Gendron-Maguire, M.; Swiatek, P.J.; Jenkins, N.A.; Copeland, N.G.; Grid
Genomics 15, 259-264, 1993
A:Title: Cloning, analysis, and chromosomal localization of Notch-1, a mouse homolog of
A:Reference number: A46019; MUID:93194170
A:Accession: A46019
A:Status: not compared with conceptual translation
A:Molecule type: nucleic acid
A:Residues: 1-2531
A:Cross-references: GB:Z11886; GB:S47228; NID:g288502; PIDN:CAA77941.1; PID:g288503
A:Note: sequence extracted from NCBI backbone (NCBIP:127318)
R:Franco del Amo, F.; Smith, D.E.; Swiatek, P.J.; Gendron-Maguire, M.; Greenspan, R.J.;
submitted to the EMBL Data Library, April 1992
A:Description: Expression pattern of Notch, a mouse homolog of Drosophila Notch, suggest
A:Reference number: S25144
A:Accession: S25144
A:Molecule type: mRNA
A:Residues: 1551-2108, 'Q', 2110-2114, 'ALP', 2118-2170 <FRA>
A:Cross-references: EMBL:Z11886

C:Genetics:
A:Gene: notch-1
A:Map position: 2
A:Note: proximal region of chromosome 2
C:Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homol
F:106-138/Domain: EGF homology <EGF1>
F:144-175/Domain: EGF homology <EGF2>
F:222-254/Domain: EGF homology <EGF3>
F:261-292/Domain: EGF homology <EGF4>
F:339-370/Domain: EGF homology <EGF5>
F:416-449/Domain: EGF homology <EGF6>
F:456-487/Domain: EGF homology <EGF7>
F:494-525/Domain: EGF homology <EGF8>
F:532-563/Domain: EGF homology <EGF9>
F:607-638/Domain: EGF homology <EGF10>
F:682-713/Domain: EGF homology <EGF11>
F:757-788/Domain: EGF homology <EGF12>
F:795-826/Domain: EGF homology <EGF13>
F:873-904/Domain: EGF homology <EGF14>
F:911-942/Domain: EGF homology <EGF15>
F:949-980/Domain: EGF homology <EGF16>
F:987-1018/Domain: EGF homology <EGF17>
F:1025-1056/Domain: EGF homology <EGF18>
F:1063-1094/Domain: EGF homology <EGF19>
F:1149-1180/Domain: EGF homology <EGF20>
F:1187-1218/Domain: EGF homology <EGF21>
F:1233-1264/Domain: EGF homology <EGF22>
F:1332-1383/Domain: EGF homology <EGF23>
F:1391-1425/Domain: EGF homology <EGF24>
F:1917-1948/Domain: ankyrin repeat homology <AN1>
F:1949-1981/Domain: ankyrin repeat homology <AN2>
F:1983-2015/Domain: ankyrin repeat homology <AN3>
F:2016-2048/Domain: ankyrin repeat homology <AN4>
F:2049-2081/Domain: ankyrin repeat homology <AN5>
Query Match 4.1%; Score 9; DB 2; Length 2531;
Best Local Similarity 100.0%; Pred. No. 0.77;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 82 PCHNGTGC 90
|||||
DB 686 PCHNGTGC 694
RESULT 8
F:75371
hypothetical protein - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
C:Accession: F75371
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.
M.; Shen, M.; Vanathavan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.;
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896
A:Accession: F75371
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-63 <WHI>
A:Cross-references: GB:AE002007; GB:AE000513; NID:g6459402; PIDN:AAF11201.1; PID:g645
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR1640
A:Map position: 1
Query Match 3.6%; Score 8; DB 2; Length 63;
Best Local Similarity 100.0%; Pred. No. 0.39;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7 AWLLVGLS 14

Db 44 AWLLVGLS 51
|||||

RESULT 9

E75609
amino acid ABC transporter, permease protein - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
C:Accession: E75609
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896
A:Accession: E75609
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-247 <WHI>
A:Cross-references: GB:AE001862; GB:AE001825; NID:g6460468; PIDN:AAF12222.1; PID:g646051
A:Experimental source: strain R1
C:Genetics:
A:Gene: DRA0138
A:Map position: 2
C:Superfamily: glycine betaine/carnitine/choline ABC transporter

Query Match 3.6%; Score 8; DB 2; Length 247;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 LLVGLSLG 16
|||||

Db 67 LLVGLSLG 74
|||||

RESULT 10

T46256
brevican - human (fragment)
N:Alternate names: protein DKFPz761L191.1
C:Species: Homo sapiens (man)
C>Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 02-Sep-2000
C:Accession: T46256
R:Ottenwälder, B.; Obermaier, B.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, January 2000
A:Reference number: 223031
A:Accession: T46256
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-330 <AAA>
A:Cross-references: EMBL:AL137504
A:Experimental source: adult amygdala; clone DKFPz761L191
C:Genetics:
A:Note: DKFPz761L191.1
C:Superfamily: aggrecan; C-type lectin homology; complement factor H repeat homology; EG

Query Match 3.6%; Score 8; DB 2; Length 330;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 82 PCHNGGTC 89
|||||

Db 73 PCHNGGTC 80
|||||

RESULT 11

T48516
probable oligopeptide transporter protein - Arabidopsis thaliana
N:Alternate names: protein F15N18.160
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000

C:Accession: T48516
R:Bevan, M.; Hilbert, H.; Braun, M.; Holzer, E.; Brandt, A.; Duesterhoeft, A.; Bancro
submitted to the Protein Sequence Database, April 2000
A:Reference number: 224490
A:Accession: T48516
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-481 <BEV>
A:Cross-references: EMBL:AL163815
A:Experimental source: cultivar Columbia; BAC clone F15N18
C:Genetics:
A:Map position: 5
A:Introns: 25/1; 97/3; 233/1; 364/3
A:Note: F15N18.160

Query Match 3.6%; Score 8; DB 2; Length 481;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 WLLVGLSL 15
|||||

Db 456 WLLVGLSL 463
|||||

RESULT 12

A31246
neurogenic protein Delta precursor - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C>Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 11-Jan-2000
C:Accession: A31246
R:Kopczynski, C.C.; Alton, A.K.; Fechtel, K.; Kooh, P.J.; Muskavitch, M.A.T.
Genes Dev. 2, 1723-1735, 1988
A:Title: Delta, a Drosophila neurogenic gene, is transcriptionally complex and encode
A:Reference number: A31246; MUID:89196890
A:Accession: A31246
A:Molecule type: mRNA
A:Residues: 1-832 <KOP>
A:Cross-references: GB:Y00222
C:Genetics:
A:Gene: FlyBase:DL
A:Cross-references: FlyBase:FBgn0000463
C:Superfamily: unassigned EGF-related proteins; EGF homology
F:295-328/Domain: EGF homology <EGX1>
F:422-450/Domain: EGF homology <EGF1>
F:457-488/Domain: EGF homology <EGF>
F:533-564/Domain: EGF homology <EGF3>

Query Match 3.6%; Score 8; DB 2; Length 832;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 82 PCHNGGTC 89
|||||

Db 537 PCHNGGTC 544
|||||

RESULT 13

S19087
gene Delta protein precursor - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 11-Jan-2000
C:Accession: S19087
R:Muskavitch, M.A.T.
submitted to the EMBL Data Library, June 1991
A:Reference number: S19087
A:Accession: S19087
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-833 <MUS>
A:Cross-references: EMBL:Y00222
C:Genetics:

A:Gene: FlyBase:DL
A:Cross-references: FlyBase:FBgn0000463
C:Superfamily: unassigned EGF-related proteins; EGF homology
F:335-371/Domain: EGF homology <EGF1>
F:378-415/Domain: EGF homology <EGX1>
F:457-488/Domain: EGF homology <EGF>
F:533-564/Domain: EGF homology <EGF3>

Query Match 3.6%; Score 8; DB 2; Length 833;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 82 PCHNGGTC 89
|||||
Db 537 PCHNGGTC 544

RESULT 14
S00670
neurogenic repetitive locus delta protein precursor - fruit fly (Drosophila melanogaster)
N:Alternate names: gene DL protein
C:Species: Drosophila melanogaster
C:Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 21-Jul-2000
C:Accession: S00670; A26637
R:Vaessin, H.; Bremer, K.A.; Knust, E.; Campos-Ortega, J.A.
EMBO J. 6, 3431-3440, 1987
A:Title: The neurogenic gene Delta of Drosophila melanogaster is expressed in neurogenic
A:Reference number: S00670
A:Accession: S00670
A:Molecule type: mRNA
A:Residues: 1-880 <VAE>
A:Cross-references: EMBL:X06289; NID:g7852; PID:g7853
R:Knust, E.; Dietrich, U.; Tepass, U.; Bremer, K.A.; Weigel, D.; Vaessin, H.; Campos-Ortega
EMBO J. 6, 761-766, 1987
A:Title: EGF homologous sequences encoded in the genome of Drosophila melanogaster, and
A:Reference number: A91081; MUID:87218537
A:Accession: A26637
A:Molecule type: mRNA
A:Residues: 422-436, 'ET', 439-458, 'A', 460-489, 'T', 491-621 <KNU>
A:Cross-references: GB:X05140; NID:g7851; PIDN:CAA28786.1; PID:g929563
C:Genetics:
A:Gene: Delta; DL
A:Cross-references: FlyBase:FBgn0000463
C:Superfamily: unassigned EGF-related proteins; EGF homology
C:Keywords: transmembrane protein
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-88/Product: neurogenic repetitive locus delta protein #status predicted <MAT>
F:457-488/Domain: EGF homology <EGF1>
F:533-564/Domain: EGF homology <EGF2>

Query Match 3.6%; Score 8; DB 2; Length 880;
Best Local Similarity 100.0%; Pred. No. 3.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 82 PCHNGGTC 89
|||||
Db 537 PCHNGGTC 544

RESULT 15
S49126
brevican precursor - rat
N:Alternate names: aggrecan-like protein
C:Species: Rattus norvegicus (Norway rat)
C:Date: 01-Feb-1995 #sequence_revision 12-May-1995 #text_change 26-May-2000
C:Accession: S49126; I55457
R:Seidenbecher, C.I.; Langnase, K.; Wax, H.; Seidel, B.; Garner, C.C.; Gundelfinger, E.
submitted to the EMBL Data Library, June 1994
A:Description: Molecular cloning of a new member of the aggrecan/versican family of proteoglycans
A:Reference number: S49126
A:Accession: S49126

A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-883 <SEI>
A:Cross-references: EMBL:X79881; NID:g509396; PIDN:CAA56255.1; PID:g509397
R:Seidenbecher, I.C.; Richter, K.; Rauch, U.; Fassler, R.; Garner, C.C.; Gundelfinger
J. Biol. Chem. 270, 27206-27212, 1995
A:Title: Brevican, a Chondroitin Sulfate Proteoglycan of Rat Brain, Occurs as Secreted
A:Reference number: I55457; MUID:96070828
A:Accession: I55457
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-883 <RES>
A:Cross-references: EMBL:X79881; NID:g509396; PIDN:CAA56255.1; PID:g509397
C:Comment: For an alternative splice form, see PIR:A53308.
C:Superfamily: aggrecan; C-type lectin homology; complement factor H repeat homology;
C:Keywords: alternative splicing
F:49-138/Domain: immunoglobulin homology <IMM>
F:173-250/Domain: link protein repeat homology <LNK1>
F:271-352/Domain: link protein repeat homology <LNK2>
F:626-657/Domain: EGF homology <EGF>
F:664-784/Domain: C-type lectin homology <LCH>
F:791-847/Domain: complement factor H repeat homology <FHD>

Query Match 3.6%; Score 8; DB 2; Length 883;
Best Local Similarity 100.0%; Pred. No. 3.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 82 PCHNGGTC 89
|||||
Db 630 PCHNGGTC 637

RESULT 16
S57653
brevican precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 19-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 13-Aug-1999
C:Accession: S57653
R:Rauch, U.; Forsberg, N.; Kulbe, G.; Arnold-Ammer, I.; Faessler, R.
submitted to the EMBL Data Library, May 1995
A:Description: Cloning and sequence of mouse neurocan and brevican and their differenc
A:Reference number: S57653
A:Accession: S57653
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-883 <RAU>
A:Cross-references: EMBL:X87096; NID:g886889; PIDN:CAA60575.1; PID:g886890
C:Superfamily: aggrecan; C-type lectin homology; complement factor H repeat homology;
F:49-138/Domain: immunoglobulin homology <IMM>
F:173-250/Domain: link protein repeat homology <LNK1>
F:271-352/Domain: link protein repeat homology <LNK2>
F:626-657/Domain: EGF homology <EGF>
F:664-784/Domain: C-type lectin homology <LCH>
F:791-847/Domain: complement factor H repeat homology <FHD>

Query Match 3.6%; Score 8; DB 2; Length 883;
Best Local Similarity 100.0%; Pred. No. 3.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 82 PCHNGGTC 89
|||||
Db 630 PCHNGGTC 637

RESULT 17
A54423
brevican precursor - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 09-Sep-1994 #sequence_revision 09-Sep-1994 #text_change 21-Jan-2000
C:Accession: A54423; S41914
R:Yamada, H.; Watanabe, K.; Shimonaka, M.; Yamaguchi, Y.

R;Ushkaryov, Y.A.; Suedhof, T.C.
Proc. Natl. Acad. Sci. U.S.A. 90, 6410-6414, 1993
A;Title: Neurexin IIIalpha: extensive alternative splicing generates membrane-bound a
A;Reference number: A48216; MUID:93342001
A;Accession: B48218

A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1471 <USH>
A:Cross-references: GB:L14851
A:Accession: C48218
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1368,1372-1471 <US2>
A:Cross-references: GB:L14851
C:Genetics:
A:introns: 1372/1
C:Superfamily: unassigned EGF-related proteins; EGF homology
C:Keywords: alternative splicing; brain; cell surface component; duplication; receptor;
F:1-27/Domain: signal sequence #status predicted <SIG>
F:202-234/Domain: EGF homology <EGF>
F:651-683/Domain: EGF homology <EGF1>

Query Match 3.6%; Score 8; DB 2; Length 1471;
Best Local Similarity 100.0%; Pred. No. 5.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 PCENGGIC 37
Db 206 PCENGGIC 213
|||||

RESULT 22
I48216
neurexin III-alpha membrane-bound type 1 precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 26-May-1994 #sequence_revision 26-May-1994 #text_change 17-Nov-2000
C:Accession: I48216; A48218; G48216; H48216; C48216; D48216; E48216; F48216
R:Ushkaryov, Y.A.; Suedhof, T.C.
Proc. Natl. Acad. Sci. U.S.A. 90, 6410-6414, 1993
A:Title: Neurexin IIIalpha: extensive alternative splicing generates membrane-bound and
A:Reference number: A48216; MUID:93342001
A:Accession: I48216
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1578 <USH>
A:Cross-references: GB:L14851
A:Accession: A48218
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1216,1',1218-1368,1372-1578 <US2>
A:Cross-references: GB:L14851
A:Accession: G48216
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1371,'A',1407-1410,'KK',1413,'NFQ',1539,'CG',1542,'S',1544,'CPR',1548,'FLH'
A:Cross-references: GB:L14851
A:Accession: H48216
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1368,'ATTGTTKK',1376,'N','FQ',1539,'CG',1542,'S',1544,'CPR',1548,'FLHNF',15'
A:Cross-references: GB:L14851
A:Accession: C48216
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1371,'VLERRIILNLTNAHP',1373-1375,'QSKTC' <US5>
A:Cross-references: GB:L14851
A:Accession: D48216
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1368,1537,'L',1539,'R',1541,'IILNLTNAHPKSLQS',1558,'TC' <US6>
A:Cross-references: GB:L14851
A:Accession: E48216
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1372,'IL',1375,'KSF' <US7>
A:Cross-references: GB:L14851
A:Accession: F48216

A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1368,'DILL',1373-1374,'F' <US8>
A:Cross-references: GB:L14851
C:Genetics:
A:introns: 1372/1
C:Superfamily: unassigned EGF-related proteins; EGF homology
C:Keywords: alternative splicing; brain; cell surface component; duplication; recepto
F:1-27/Domain: signal sequence #status predicted <SIG>
F:202-234/Domain: EGF homology <EGF>
F:651-683/Domain: EGF homology <EGF1>

Query Match 3.6%; Score 8; DB 2; Length 1578;
Best Local Similarity 100.0%; Pred. No. 6.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 PCENGGIC 37
Db 206 PCENGGIC 213
|||||

RESULT 23
C40228
neurexin II-alpha precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 21-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 11-Jan-2000
C:Accession: C40228; S27886; S27887
R:Ushkaryov, Y.A.; Petrenko, A.G.; Geppert, M.; Suedhof, T.C.
Science 257, 50-56, 1992
A:Title: Neurexins: synaptic cell surface proteins related to the alpha-latrotoxin re
A:Reference number: A40228; MUID:92320296
A:Accession: C40228
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-1715 <USH>
A:Cross-references: GB:M96376; NID:g205714; PIDN:AAA41707.1; PID:g205716
A:Note: authors translated the codon GAC for residue 1464 as Thr and ACC for residue
R:Ushkaryov, Y.A.; Petrenko, A.G.; Geppert, M.; Sudhof, T.C.
submitted to the EMBL Data Library, July 1992
A:Description: Neurexins: Synaptic cell surface proteins related to the alpha-latroto
A:Reference number: S27884
A:Accession: S27886
A:Molecule type: mRNA
A:Residues: 1-1666,'CRK',1670,'PREKKLLPG',1683-1685,'GL',1688,'LDLA',1694-1695,'CCVCR'
A:Cross-references: EMBL:M96376; NID:g205714; PIDN:AAA41706.1; PID:g205715
C:Genetics:
A:introns: 1666/2
C:Superfamily: unassigned EGF-related proteins; EGF homology
C:Keywords: alternative splicing; transmembrane protein
F:1-27/Domain: signal sequence #status predicted <SIG>
F:28-1715/Product: neurexin II-alpha #status predicted <MAT>
F:694-726/Domain: EGF homology <EGF>
F:1103-1135/Domain: EGF homology <EGF1>

Query Match 3.6%; Score 8; DB 2; Length 1715;
Best Local Similarity 100.0%; Pred. No. 6.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 PCENGGIC 134
Db 698 PCENGGIC 705
|||||

RESULT 24
A35844
Xotch protein - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C:Date: 12-Oct-1990 #sequence_revision 12-Oct-1990 #text_change 13-Aug-1999
C:Accession: A35844
R:Coffman, C.; Harris, W.; Kintner, C.
Science 249, 1438-1441, 1990

A;Title: Notch, the Xenopus homolog of Drosophila notch.
A;Reference number: A35844; MUID:90385285
A;Accession: A35844
A;Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra
A;Molecule type: mRNA
A;Residues: 1-2524 <COP>

C;Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homology
F;146-177/Domain: EGF homology <EGX1>
F;184-215/Domain: EGF homology <EGF1>
F;222-254/Domain: EGF homology <EGF>
F;456-487/Domain: EGF homology <EGX2>
F;757-788/Domain: EGF homology <EGF3>
F;1025-1036/Domain: EGF homology <EGX3>
F;1924-1956/Domain: ankyrin repeat homology <AN1>
F;1957-1989/Domain: ankyrin repeat homology <AN2>
F;1991-2023/Domain: ankyrin repeat homology <AN3>
F;2024-2056/Domain: ankyrin repeat homology <AN4>
F;2057-2089/Domain: ankyrin repeat homology <AN5>

Query Match 3.6%; Score 8; DB 2; Length 2524;
Best Local Similarity 100.0%; Pred. No. 9.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 PNPCHNGG 87
|||||||
Db 913 PNPCHNGG 920

RESULT 25

T31070
notch homolog - sea urchin (Lytechinus variegatus)
C;Species: Lytechinus variegatus (variegated urchin)
C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 31-Jan-2000
C;Accession: T31070
R;Sherwood, D.R.; McClay, D.R.
Development 124, 3363-3374, 1997
A;Title: Identification and localization of a sea urchin Notch homologue: insights into
A;Reference number: Z20966; MUID:97454256
A;Accession: T31070
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-2531 <SHE>
A;Cross-references: EMBL:AF000634; MID:g2570350; PID:g2570351; PIDN:AAB82088.1
C;Superfamily: notch protein; ankyrin repeat homology; EGF homology

Query Match 3.6%; Score 8; DB 2; Length 2531;
Best Local Similarity 100.0%; Pred. No. 9.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 82 PCHNGGTC 89
|||||||
Db 673 PCHNGGTC 680

Search completed: May 23, 2001, 06:22:20
Job time: 239 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 06:25:03 ; Search time 17.36 Seconds
(without alignments)
436.087 Million cell updates/sec

Title: US-09-237-981-29
Perfect score: 221
Sequence: 1 MKHLVAAMLLVGLSLGVPQF.....INAWTAENDRWPIQVTG 221

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 93435 seqs, 34255486 residues

Word size : 0

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 50 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	10	4.5	2444	1 NTC1_HUMAN	P46531 homo sapien
2	9	4.1	1408	1 SERD_DROME	P18168 drosophila
3	9	4.1	2139	1 CREB_DROME	P10040 drosophila
4	9	4.1	2437	1 NOTC_BRARE	P46530 brachydanio
5	9	4.1	2531	1 NTC1_MOUSE	Q01708 mus musculus
6	9	4.1	2531	1 NTC1_MOUSE	Q07008 rattus norv
7	8	3.6	833	1 DL_DROME	P10041 drosophila
8	8	3.6	883	1 PGCB_MOUSE	Q61361 mus musculus
9	8	3.6	883	1 PGCB_RAT	P55068 rattus norv
10	8	3.6	912	1 PGCB_BOVIN	Q28062 bos taurus
11	8	3.6	2524	1 NOTC_XENLA	P21783 xenopus lae
12	8	3.6	2670	1 YAO5_SCHPO	Q10105 schizosacch
13	7	3.2	106	1 SH_RAT	P55248 rattus norv
14	7	3.2	110	1 CYOD_PSEPU	Q9wwr4 pseudomonas
15	7	3.2	149	1 CRAA_RANTE	P02508 rana escule
16	7	3.2	167	1 CRAA_RANES	P02507 rana escule
17	7	3.2	173	1 CRAA_ALLMI	P06904 alligator m
18	7	3.2	173	1 CRAA_CHICK	P02489 homo sapien
19	7	3.2	173	1 CRAA_HUMAN	Q91311 rana catesb
20	7	3.2	173	1 CRAA_RANCA	P02505 rhea american
21	7	3.2	173	1 CRAA_RHEAM	P02506 tupinambis
22	7	3.2	259	1 YL23_YEAST	P53769 saccharomyc
23	7	3.2	308	1 MENA_ECOLI	P32166 escherichia
24	7	3.2	349	1 CGN6_HUMAN	P31997 homo sapien
25	7	3.2	358	1 GPRK_HUMAN	Q99678 homo sapien
26	7	3.2	387	1 MFGM_HUMAN	Q08431 homo sapien
27	7	3.2	402	1 UDG_STRPY	Q07172 streptococc
28	7	3.2	450	1 CREB_ECOLI	P08369 escherichia
29	7	3.2	562	1 TPA_HUMAN	P00750 homo sapien
30	7	3.2	575	1 TREM_HUMAN	P07204 homo sapien
31	7	3.2	577	1 TRBM_MOUSE	P15306 mus musculus
32	7	3.2	586	1 COS_FUGRU	P79755 fugu rubrip
33	7	3.2			

RESULT 1

ID	NTC1_HUMAN	STANDARD;	PRT;	2444 AA.
AC	P46531;			
DT	01-NOV-1995 (Rel. 32, Created)			
DT	01-NOV-1995 (Rel. 32, Last sequence update)			
DT	01-FEB-1996 (Rel. 33, Last annotation update)			
DE	NEUROGENIC LOCUS NOTCH PROTEIN HOMOLOG 1 PRECURSOR (TRANSLATION-			
DE	ASSOCIATED NOTCH PROTEIN TAN-1) (FRAGMENT).			
GN	NOTCH1 OR TAN1.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=91347367; PubMed=1831692;			
RA	Ellisen L.W., Bird J., West D.C., Soreng A.L., Reynolds T.C.,			
RA	Smith S.D., Sklar J.;			
RT	"TAN-1, the human homolog of the Drosophila notch gene, is broken by			
RT	chromosomal translocations in T lymphoblastic neoplasms.";			
RL	Cell 66:649-661(1991).			
CC	!- FUNCTION: MAY BE IMPORTANT FOR NORMAL LYMPHOCYTE FUNCTION. IN			
CC	ALTERED FORM, MAY CONTRIBUTE TO TRANSFORMATION OR PROGRESSION			
CC	IN SOME T-CELL NEOPLASMS.			
CC	!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.			
CC	!- TISSUE SPECIFICITY: IN FETAL TISSUES MOST ABUNDANT IN SPLEEN,			
CC	BRAIN STEM AND LUNG. ALSO PRESENT IN MOST ADULT TISSUES WHERE IT			
CC	IS FOUND MAINLY IN LYMPHOID TISSUES.			
CC	!- SIMILARITY: HIGH, WITH OTHER NOTCH-TYPE PROTEINS.			
CC	!- SIMILARITY: CONTAINS 36 EGF-LIKE DOMAINS.			
CC	!- SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.			
CC	!- SIMILARITY: CONTAINS 5 ANK REPEATS.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
CC	EMBL; M73980; AAA60614.1; --			
DR	HSSP; P00740; IIXA.			
DR	MIN; 190198; --			
DR	InterPro; IPR000152; --			
DR	InterPro; IPR000561; --			
DR	InterPro; IPR000800; --			
DR	InterPro; IPR001881; --			
DR	InterPro; IPR002110; --			
DR	Pfam; PF000008; EGF; 36.			

ALIGNMENTS

34	7	3.2	745	1	ZP2_HUMAN	Q05996 homo sapien
35	7	3.2	938	1	NMZ1_HUMAN	Q05586 homo sapien
36	7	3.2	938	1	NMZ1_MOUSE	P35438 mus musculus
37	7	3.2	938	1	NMZ1_RAT	P35439 rattus norv
38	7	3.2	1238	1	B3A2_CAVPO	Q92058 cavia porce
39	7	3.2	1247	1	POLS_ONNVG	P22056 o'nyong-nyo
40	7	3.2	1557	1	DVAL_DICVI	Q24702 dictyocaulu
41	7	3.2	2318	1	NTC3_MOUSE	Q61982 mus musculus
42	6	2.7	18	1	YPE4_LACLC	P42022 lactococcus
43	6	2.7	38	1	BD01_BOVIN	P46159 bos taurus
44	6	2.7	48	1	HXB6_XENLA	P31256 xenopus lae
45	6	2.7	52	1	CTL2_NPVOP	O10286 orgyia pseu
46	6	2.7	94	1	MBC_ECOLI	P07114 escherichia
47	6	2.7	120	1	YQJU_BACSU	P54558 bacillus su
48	6	2.7	124	1	UL39_HCMVA	P16813 human cytom
49	6	2.7	142	1	MK_CHICK	P24052 gallus gall
50	6	2.7	148	1	PER1_MESCR	O04683 mesembryant

DR	Pfam: PF00023; ank; 6.	FT	DISULFID	106	117	BY SIMILARITY.
DR	Pfam: PF00066; notch; 3.	FT	DISULFID	111	127	BY SIMILARITY.
DR	PROSITE; PS50088; ANK_REPEAT; 4.	FT	DISULFID	129	138	BY SIMILARITY.
DR	PROSITE; PS50297; ANK_REPEAT; 1.	FT	DISULFID	144	155	BY SIMILARITY.
DR	PROSITE; PS00010; ASK_HYDROXYL; 20.	FT	DISULFID	149	164	BY SIMILARITY.
DR	PROSITE; PS00022; EGF_1; 34.	FT	DISULFID	166	175	BY SIMILARITY.
DR	PROSITE; PS01186; EGF_2; 26.	FT	DISULFID	182	195	BY SIMILARITY.
DR	PROSITE; PS01187; EGF_CA; 18.	FT	DISULFID	189	204	BY SIMILARITY.
KW	Differentiation; Neurogenesis; Repeat; ANK repeat; EGF-like domain;	FT	DISULFID	206	215	BY SIMILARITY.
KW	Transmembrane; Signal; Glycoprotein.	FT	DISULFID	222	233	BY SIMILARITY.
FT	SIGNAL 1 18	FT	DISULFID	227	243	BY SIMILARITY.
FT	CHAIN 19 >2444	FT	DISULFID	227	243	BY SIMILARITY.
FT	DOMAIN 19 1736	FT	DISULFID	245	254	BY SIMILARITY.
FT	TRANSMEM 1737 1757	FT	DISULFID	261	272	BY SIMILARITY.
FT	DOMAIN 1758 >2444	FT	DISULFID	266	281	BY SIMILARITY.
FT	DOMAIN 20 58	FT	DISULFID	283	292	BY SIMILARITY.
FT	DOMAIN 59 99	FT	DISULFID	299	312	BY SIMILARITY.
FT	DOMAIN 102 139	FT	DISULFID	306	321	BY SIMILARITY.
FT	DOMAIN 140 176	FT	DISULFID	323	332	BY SIMILARITY.
FT	DOMAIN 178 216	FT	DISULFID	339	350	BY SIMILARITY.
FT	DOMAIN 218 255	FT	DISULFID	344	359	BY SIMILARITY.
FT	DOMAIN 257 293	FT	DISULFID	361	370	BY SIMILARITY.
FT	DOMAIN 295 333	FT	DISULFID	376	387	BY SIMILARITY.
FT	DOMAIN 335 371	FT	DISULFID	381	398	BY SIMILARITY.
FT	DOMAIN 372 410	FT	DISULFID	400	409	BY SIMILARITY.
FT	DOMAIN 412 450	FT	DISULFID	416	429	BY SIMILARITY.
FT	DOMAIN 452 488	FT	DISULFID	423	438	BY SIMILARITY.
FT	DOMAIN 490 526	FT	DISULFID	440	449	BY SIMILARITY.
FT	DOMAIN 528 564	FT	DISULFID	456	467	BY SIMILARITY.
FT	DOMAIN 566 601	FT	DISULFID	461	476	BY SIMILARITY.
FT	DOMAIN 603 639	FT	DISULFID	478	487	BY SIMILARITY.
FT	DOMAIN 641 676	FT	DISULFID	494	505	BY SIMILARITY.
FT	DOMAIN 678 714	FT	DISULFID	499	514	BY SIMILARITY.
FT	DOMAIN 716 751	FT	DISULFID	516	525	BY SIMILARITY.
FT	DOMAIN 753 789	FT	DISULFID	532	543	BY SIMILARITY.
FT	DOMAIN 829 868	FT	DISULFID	537	552	BY SIMILARITY.
FT	DOMAIN 906 944	FT	DISULFID	554	563	BY SIMILARITY.
FT	DOMAIN 946 982	FT	DISULFID	570	580	BY SIMILARITY.
FT	DOMAIN 1020 1058	FT	DISULFID	575	589	BY SIMILARITY.
FT	DOMAIN 1060 1096	FT	DISULFID	591	600	BY SIMILARITY.
FT	DOMAIN 1098 1144	FT	DISULFID	607	618	BY SIMILARITY.
FT	DOMAIN 1146 1182	FT	DISULFID	612	627	BY SIMILARITY.
FT	DOMAIN 1184 1220	FT	DISULFID	629	638	BY SIMILARITY.
FT	DOMAIN 1222 1266	FT	DISULFID	645	655	BY SIMILARITY.
FT	DOMAIN 1268 1306	FT	DISULFID	650	664	BY SIMILARITY.
FT	DOMAIN 1308 1347	FT	DISULFID	666	675	BY SIMILARITY.
FT	DOMAIN 1349 1385	FT	DISULFID	682	693	BY SIMILARITY.
FT	DOMAIN 1427 1481	FT	DISULFID	687	702	BY SIMILARITY.
FT	REPEAT 1446 1481	FT	DISULFID	704	713	BY SIMILARITY.
FT	REPEAT 1482 1523	FT	DISULFID	720	730	BY SIMILARITY.
FT	REPEAT 1524 1563	FT	DISULFID	725	739	BY SIMILARITY.
FT	REPEAT 1524 1563	FT	DISULFID	741	750	BY SIMILARITY.
FT	REPEAT 1524 1563	FT	DISULFID	757	768	BY SIMILARITY.
FT	REPEAT 1524 1563	FT	DISULFID	762	777	BY SIMILARITY.
FT	REPEAT 1524 1563	FT	DISULFID	779	788	BY SIMILARITY.
FT	REPEAT 1524 1563	FT	DISULFID	795	806	BY SIMILARITY.
FT	REPEAT 1524 1563	FT	DISULFID	800	815	BY SIMILARITY.
FT	REPEAT 1524 1563	FT	DISULFID	817	826	BY SIMILARITY.
FT	REPEAT 1524 1563	FT	DISULFID	833	844	BY SIMILARITY.
FT	REPEAT 1524 1563	FT	DISULFID	838	855	BY SIMILARITY.
FT	REPEAT 1524 1563	FT	DISULFID	857	867	BY SIMILARITY.
FT	REPEAT 1524 1563	FT	DISULFID	874	885	BY SIMILARITY.
FT	REPEAT 1524 1563	FT	DISULFID	879	894	BY SIMILARITY.
FT	REPEAT 1524 1563	FT	DISULFID	896	905	BY SIMILARITY.
FT	REPEAT 1524 1563	FT	DISULFID	912	923	BY SIMILARITY.
FT	REPEAT 1524 1563	FT	DISULFID	917	932	BY SIMILARITY.
FT	REPEAT 1524 1563	FT	DISULFID	934	943	BY SIMILARITY.
FT	REPEAT 1524 1563	FT	DISULFID	948	959	BY SIMILARITY.
FT	REPEAT 1524 1563	FT	DISULFID	988	999	BY SIMILARITY.
FT	REPEAT 1524 1563	FT	DISULFID	993	1008	BY SIMILARITY.
FT	REPEAT 1524 1563	FT	DISULFID	1010	1019	BY SIMILARITY.
FT	REPEAT 1524 1563	FT	DISULFID	1026	1037	BY SIMILARITY.
FT	REPEAT 1524 1563	FT	DISULFID	1031	1046	BY SIMILARITY.
FT	REPEAT 1524 1563	FT	DISULFID	1048	1057	BY SIMILARITY.
FT	REPEAT 1524 1563	FT	DISULFID	1064	1075	BY SIMILARITY.

FT DISULFID 1069 1084 BY SIMILARITY.
FT DISULFID 1086 1095 BY SIMILARITY.
FT DISULFID 1102 1123 BY SIMILARITY.
FT DISULFID 1117 1132 BY SIMILARITY.
FT DISULFID 1134 1143 BY SIMILARITY.

Query Match 4.5%; Score 10; DB 1; Length 2444;
Best Local Similarity 100.0%; Pred. NO. 0.048;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 NPCHNGTCE 90
DB 685 NPCHNGTCE 694

RESULT 2
SERR_DROME STANDARD; PRT: 1408 AA.
AC P18168;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE SERRATE PROTEIN PRECURSOR (BEADED PROTEIN).
GN SER OR BD.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_taxid=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=91347903; PubMed=1840519;
RA Thomas U., Speicher S.A., Knust E.;
RT "The Drosophila gene Serrate encodes an EGF-like transmembrane protein with a complex expression pattern in embryos and wing discs.";
RT Development 111:749-761(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=91099666; PubMed=2125287;
RA Fleming R.J., Scottgale T.N., Diederich R.J., Artavanis-Tsakonas S.;
RT "The gene Serrate encodes a putative EGF-like transmembrane protein essential for proper ectodermal development in Drosophila melanogaster.";
RT Genes Dev. 4:2188-2201(1990).
CC -!- FUNCTION: ESSENTIAL FOR PROPER ECTODERMAL DEVELOPMENT. SERRATE MAY REPRESENT AN ELEMENT IN A NETWORK OF INTERACTING MOLECULES OPERATING AT THE CELL SURFACE DURING THE DIFFERENTIATION OF CERTAIN TISSUES.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- TISSUE SPECIFICITY: APPEARS TO BE RESTRICTED EXCLUSIVELY TO CELLS OF ECTODERMAL ORIGIN.
CC -!- MISCELLANEOUS: SEPARATION OF NEUROBLASTS FROM THE ECTODERM INTO THE INNER PART OF EMBRYO IS ONE OF THE FIRST STEPS OF CNS DEVELOPMENT IN INSECTS, THIS PROCESS IS UNDER CONTROL OF THE NEUROGENIC GENES.
CC -!- MISCELLANEOUS: NOTCH AND SERRATE MAY INTERACT AT THE PROTEIN LEVEL, IT IS CONCEIVABLE THAT THE SERRATE AND DELTA PROTEINS MAY COMPETE FOR BINDING WITH THE NOTCH PROTEIN.
CC -!- SIMILARITY: CONTAINS 14 EGF-LIKE DOMAINS.
CC -!- SIMILARITY: BELONGS TO THE DELTA/SERRATE/JAGGED FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; X56811; CAA40148.1; -;
CC EMBL; M35759; AAA28938.1; -;

DR PIR: A36666; A36666.
DR PIR: S16878; S16878.
DR HSP; P00743; LWHE.
DR FlyBase; FBgn0004197; Ser.
DR InterPro; IPR000152; -
DR InterPro; IPR000561; -
DR InterPro; IPR001438; -
DR InterPro; IPR001774; -
DR InterPro; IPR001881; -
DR Pfam; PF01414; DSL; 1.
DR Pfam; PF00008; EGF; 11.
DR PRINTS; PR00010; EGFBLD.
DR PROSITE; PS00010; ASX HYDROXYL; 7.
DR PROSITE; PS00022; EGF_1; 14.
DR PROSITE; PS01186; EGF_2; 8.
DR PROSITE; PS01187; EGF_CA; 5.
KW Differentiation; Repeat; EGF-like domain; Transmembrane;
KW Glycoprotein; Signal.
FT SIGNAL 1 83 POTENTIAL.
FT CHAIN 84 1408 SERRATE PROTEIN.
FT DOMAIN 84 1223 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1224 1249 POTENTIAL.
FT DOMAIN 1250 1408 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 284 317 EGF-LIKE 1.
FT DOMAIN 315 349 EGF-LIKE 2.
FT DOMAIN 351 389 EGF-LIKE 3.
FT DOMAIN 391 489 EGF-LIKE 4.
FT DOMAIN 407 476 SER-RICH (INSERT).
FT DOMAIN 491 527 EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 529 609 EGF-LIKE 6, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 611 646 EGF-LIKE 7, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 648 684 EGF-LIKE 8, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 686 721 EGF-LIKE 9.
FT DOMAIN 723 797 EGF-LIKE 10.
FT DOMAIN 737 769 THR-RICH (INSERT).
FT DOMAIN 799 835 EGF-LIKE 11, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 837 877 EGF-LIKE 12.
FT DOMAIN 879 915 EGF-LIKE 13.
FT DOMAIN 917 953 EGF-LIKE 14, CALCIUM-BINDING (POTENTIAL).
FT DISULFID 288 299 BY SIMILARITY.
FT DISULFID 292 305 BY SIMILARITY.
FT DISULFID 307 316 BY SIMILARITY.
FT DISULFID 319 330 BY SIMILARITY.
FT DISULFID 325 337 BY SIMILARITY.
FT DISULFID 339 348 BY SIMILARITY.
FT DISULFID 355 367 BY SIMILARITY.
FT DISULFID 361 377 BY SIMILARITY.
FT DISULFID 379 388 BY SIMILARITY.
FT DISULFID 395 406 BY SIMILARITY.
FT DISULFID 400 477 BY SIMILARITY.
FT DISULFID 479 488 BY SIMILARITY.
FT DISULFID 495 506 BY SIMILARITY.
FT DISULFID 500 515 BY SIMILARITY.
FT DISULFID 517 526 BY SIMILARITY.
FT DISULFID 533 588 BY SIMILARITY.
FT DISULFID 582 597 BY SIMILARITY.
FT DISULFID 599 608 BY SIMILARITY.
FT DISULFID 615 625 BY SIMILARITY.
FT DISULFID 619 634 BY SIMILARITY.
FT DISULFID 636 645 BY SIMILARITY.
FT DISULFID 652 663 BY SIMILARITY.
FT DISULFID 657 672 BY SIMILARITY.
FT DISULFID 674 683 BY SIMILARITY.
FT DISULFID 690 700 BY SIMILARITY.
FT DISULFID 695 709 BY SIMILARITY.
FT DISULFID 711 720 BY SIMILARITY.
FT DISULFID 803 814 BY SIMILARITY.
FT DISULFID 808 823 BY SIMILARITY.
FT DISULFID 825 834 BY SIMILARITY.
FT DISULFID 841 852 BY SIMILARITY.
FT DISULFID 846 865 BY SIMILARITY.
FT DISULFID 867 876 BY SIMILARITY.
FT DISULFID 883 894 BY SIMILARITY.

```

FT DISULFID 888 903 BY SIMILARITY.
FT DISULFID 905 914 BY SIMILARITY.
FT DISULFID 921 932 BY SIMILARITY.
FT DISULFID 926 941 BY SIMILARITY.
FT DISULFID 943 952 BY SIMILARITY.
FT CARBOHYD 152 152 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 152 152 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 247 247 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 331 331 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 412 412 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 452 452 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 558 558 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 739 739 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 965 965 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 977 977 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1004 1004 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1030 1030 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1150 1150 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 14 17 MISSING (IN REF. 2).
FT CONFLICT 27 27 P -> A (IN REF. 2).
FT CONFLICT 1352 1352 T -> S (IN REF. 2).
SQ SEQUENCE 1408 AA; 150660 MW; 569DA4270A9C7840 CRC64;

Query Match 4.1%; Score 9; DB 1; Length 1408;
Best Local Similarity 100.0%; Pred. No. 0.34;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 82 PCHNGTCE 90
DB 887 PCHNGTCE 895

RESULT 3
CRB_DROME STANDARD; PRT; 2139 AA.
AC P10040;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE CRUMBS PROTEIN PRECURSOR (95F).
GN CRB
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OREGON-R; TISSUE=Embryo;
RX MEDLINE=90263104; PubMed=2344615;
RA Tepass U., Theres C., Knust E.;
RT "Crumbs encodes an EGF-like protein expressed on apical membranes of
RT Drosophila epithelial cells and required for organization of
RT epithelia.";
RL Cell 61:787-799(1990).
RN [2]
RP SEQUENCE OF 1663-1955 FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=87218537; PubMed=3107986;
RA Knust E., Dietrich U., Tepass U., Bremer K.A., Weigel D.,
RA Vaessin H., Campos-Ortega J.A.;
RT "Egf homologous sequences encoded in the genome of Drosophila
RT melanogaster, and their relation to neurogenic genes.";
RL EMBO J. 6:761-766(1987).
CC -1- FUNCTION: MAY PLAY A ROLE IN THE DEVELOPMENT OF EPITHELIA,
CC POSSIBLY FOR THE ESTABLISHMENT AND/OR MAINTENANCE OF CELL
CC POLARITY. IT MAY ACT AS A SIGNAL.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- PTM: PHOSPHORYLATED IN THE CYTOPLASMIC DOMAIN (POTENTIAL).
CC -1- SIMILARITY: CONTAINS 29 EGF-LIKE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL; M33753; CAA28428.1; ALT_SEQ.
EMBL; X05144; CAA28793.1;
PIR; B26637; B26637.
PIR; A35672; A35672.
HSP; P00740; IIXA.
DR EMBL; X05144; CAA28793.1;
DR PIR; B26637; B26637.
DR PIR; A35672; A35672.
DR HSP; P00740; IIXA.
DR FlyBase; FBgn000368; crb.
DR InterPro; IPR000152;
DR InterPro; IPR000561;
DR InterPro; IPR001438;
DR InterPro; IPR001791;
DR InterPro; IPR001881;
DR Pfam; PF00008; EGF; 27;
DR Pfam; PF00054; laminin_G; 3.
DR PRINTS; PRO0010; EGFBL00.
DR PROSITE; PS00010; ASX_HYDROXYL; 15.
DR PROSITE; PS00022; EGF_1; 26.
DR PROSITE; PS01186; EGF_2; 17.
DR PROSITE; PS01187; EGF_CA; 15.
KW Differentiation; Repeat; EGF-like domain; Transmembrane;
KW Glycoprotein; Signal; Phosphorylation.
FT SIGNAL 1 90
FT CHAIN 91 2139 CRUMBS PROTEIN.
FT DOMAIN 91 2084 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 2085 2111 POTENTIAL.
FT DOMAIN 2112 2139 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 267 303 EGF-LIKE 1.
FT DOMAIN 306 343 EGF-LIKE 2.
FT DOMAIN 348 386 EGF-LIKE 3.
FT DOMAIN 388 425 EGF-LIKE 4.
FT DOMAIN 427 463 EGF-LIKE 5.
FT DOMAIN 464 500 EGF-LIKE 6.
FT DOMAIN 501 532 EGF-LIKE 7.
FT DOMAIN 545 581 EGF-LIKE 8.
FT DOMAIN 582 611 EGF-LIKE 9.
FT DOMAIN 609 645 EGF-LIKE 10.
FT DOMAIN 648 685 EGF-LIKE 11.
FT DOMAIN 687 723 EGF-LIKE 12.
FT DOMAIN 725 761 EGF-LIKE 13.
FT DOMAIN 763 800 EGF-LIKE 14.
FT DOMAIN 802 838 EGF-LIKE 15.
FT DOMAIN 840 902 EGF-LIKE 16.
FT DOMAIN 904 940 EGF-LIKE 17.
FT DOMAIN 942 978 EGF-LIKE 18.
FT DOMAIN 980 1021 EGF-LIKE 19.
FT DOMAIN 1207 1243 EGF-LIKE 20.
FT DOMAIN 1481 1517 EGF-LIKE 21.
FT DOMAIN 1759 1795 EGF-LIKE 22.
FT DOMAIN 1797 1833 EGF-LIKE 23.
FT DOMAIN 1835 1871 EGF-LIKE 24.
FT DOMAIN 1874 1915 EGF-LIKE 25.
FT DOMAIN 1915 1951 EGF-LIKE 26.
FT DOMAIN 1953 1989 EGF-LIKE 27.
FT DOMAIN 1991 2029 EGF-LIKE 28.
FT DOMAIN 2030 2070 EGF-LIKE 29.
FT DISULFID 271 282 BY SIMILARITY.
FT DISULFID 276 291 BY SIMILARITY.
FT DISULFID 293 302 BY SIMILARITY.
FT DISULFID 310 321 BY SIMILARITY.
FT DISULFID 315 331 BY SIMILARITY.
FT DISULFID 333 342 BY SIMILARITY.
FT DISULFID 352 363 BY SIMILARITY.
FT DISULFID 357 374 BY SIMILARITY.
FT DISULFID 376 385 BY SIMILARITY.
FT DISULFID 392 403 BY SIMILARITY.
FT DISULFID 397 412 BY SIMILARITY.
FT DISULFID 414 424 BY SIMILARITY.

```



```
FT DISULFID 431 442 BY SIMILARITY.
FT DISULFID 436 451 BY SIMILARITY.
FT DISULFID 453 462 BY SIMILARITY.
FT DISULFID 468 479 BY SIMILARITY.
FT DISULFID 473 488 BY SIMILARITY.
FT DISULFID 490 499 BY SIMILARITY.
FT DISULFID 505 515 BY SIMILARITY.
FT DISULFID 509 520 BY SIMILARITY.
FT DISULFID 522 531 BY SIMILARITY.
FT DISULFID 549 562 BY SIMILARITY.
FT DISULFID 556 569 BY SIMILARITY.
FT DISULFID 571 580 BY SIMILARITY.
FT DISULFID 586 597 BY SIMILARITY.
FT DISULFID 591 602 BY SIMILARITY.
FT DISULFID 604 610 BY SIMILARITY.
FT DISULFID 613 624 BY SIMILARITY.
FT DISULFID 618 634 BY SIMILARITY.
FT DISULFID 636 645 BY SIMILARITY.
FT DISULFID 652 664 BY SIMILARITY.
FT DISULFID 659 673 BY SIMILARITY.
FT DISULFID 675 684 BY SIMILARITY.
FT DISULFID 691 702 BY SIMILARITY.
FT DISULFID 696 711 BY SIMILARITY.
FT DISULFID 713 722 BY SIMILARITY.
FT DISULFID 729 740 BY SIMILARITY.
FT DISULFID 734 749 BY SIMILARITY.
FT DISULFID 751 760 BY SIMILARITY.
FT DISULFID 767 778 BY SIMILARITY.
FT DISULFID 772 787 BY SIMILARITY.
FT DISULFID 789 799 BY SIMILARITY.
FT DISULFID 806 817 BY SIMILARITY.
FT DISULFID 811 826 BY SIMILARITY.
FT DISULFID 828 837 BY SIMILARITY.
FT DISULFID 844 855 BY SIMILARITY.
FT DISULFID 849 890 BY SIMILARITY.
FT DISULFID 892 901 BY SIMILARITY.
FT DISULFID 908 919 BY SIMILARITY.
FT DISULFID 930 928 BY SIMILARITY.
FT DISULFID 933 939 BY SIMILARITY.
FT DISULFID 946 957 BY SIMILARITY.
FT DISULFID 952 966 BY SIMILARITY.
FT DISULFID 968 977 BY SIMILARITY.
FT DISULFID 984 995 BY SIMILARITY.
FT DISULFID 989 1009 BY SIMILARITY.
FT DISULFID 1011 1020 BY SIMILARITY.
FT DISULFID 1211 1222 BY SIMILARITY.
FT DISULFID 1216 1231 BY SIMILARITY.
FT DISULFID 1233 1242 BY SIMILARITY.
FT DISULFID 1485 1496 BY SIMILARITY.
FT DISULFID 1490 1505 BY SIMILARITY.
FT DISULFID 1507 1516 BY SIMILARITY.
FT DISULFID 1763 1774 BY SIMILARITY.
FT DISULFID 1768 1783 BY SIMILARITY.
FT DISULFID 1785 1794 BY SIMILARITY.
FT DISULFID 1801 1812 BY SIMILARITY.
FT DISULFID 1806 1821 BY SIMILARITY.
FT DISULFID 1823 1832 BY SIMILARITY.
FT DISULFID 1839 1850 BY SIMILARITY.
FT DISULFID 1844 1859 BY SIMILARITY.
FT DISULFID 1861 1870 BY SIMILARITY.
FT DISULFID 1878 1889 BY SIMILARITY.
FT DISULFID 1883 1903 BY SIMILARITY.
FT DISULFID 1905 1914 BY SIMILARITY.
FT DISULFID 1919 1930 BY SIMILARITY.
FT DISULFID 1924 1939 BY SIMILARITY.
FT DISULFID 1941 1950 BY SIMILARITY.
FT DISULFID 1957 1968 BY SIMILARITY.
FT DISULFID 1962 1977 BY SIMILARITY.
FT DISULFID 1979 1988 BY SIMILARITY.
FT DISULFID 1995 2008 BY SIMILARITY.
FT DISULFID 2002 2017 BY SIMILARITY.
FT DISULFID 2019 2028 BY SIMILARITY.
FT CARBOHYD 37 N-LINKED (GLCNAC. . .) (POTENTIAL).

Query Match 4.1%; Score 9; DB 1; Length 2139;
Best Local Similarity 100.0%; Pred. No. 0.48;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 48 CECPEGFAG 56
      |||||
DB 2058 CECPEGFAG 2066

RESULT 4
NOTC BRARE
ID NOTC BRARE STANDARD; PRT; 2437 AA.
AC P46530;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE NEUROGENIC LOCUS NOTCH HOMOLOG PROTEIN PRECURSOR.
GN NOTCH.
OS Brachydanio rerio (Zebrafish) (Zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Rasbora; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=94128602; PubMed=8297791;
RA Bierkamp C., Campos-Ortega J.A.;
RA "A zebrafish homologue of the Drosophila neurogenic gene Notch and
RL its pattern of transcription during early embryogenesis.";
RL Mech. Dev. 43:87-100(1993).
CC -1- FUNCTION: IMPLICATED IN CELL FATE SPECIFICATIONS DURING THE
CC EMBRYO DEVELOPMENT. MAY BE INVOLVED IN THE FORMATION OF THE
CC NEURAL PLATE, NOTOCHORD AND BRAIN VESICLES.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN ALL CELLS IN PREGASTRULATION
CC STAGES. DURING GASTRULATION IS DIFFERENTIALLY EXPRESSED,
CC ACCUMULATING PREDOMINANTLY IN THE PRECHORDAL MESODERM AND
CC NOTOCHORD. AT THE END OF GASTRULATION, EXPRESSED ALONG THE
CC ANTERIOR-POSTERIOR AXIS INCLUDING THE DEVELOPING NEURAL PLATE
CC AND DIFFERENTIATING MESODERM. ALSO PRESENT IN THE DEVELOPING
CC BRAIN AND HEAD REGIONS.
CC -1- SIMILARITY: HIGH, WITH OTHER NOTCH-TYPE PROTEINS.
CC -1- SIMILARITY: CONTAINS 36 EGF-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
CC -1- SIMILARITY: CONTAINS 6 ANK REPEATS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DB EMBL; X69088; CAA48031.1; -
```

FT	DOMAIN	2265	2276	POLY-GLN (OPA-REPEAT) :
FT	DOMAIN	2265	2276	POLY-GLN (OPA-REPEAT) :

```

FT DISULFID 952 967 BY SIMILARITY.
FT DISULFID 969 978 BY SIMILARITY.
FT DISULFID 1023 1034 BY SIMILARITY.
FT DISULFID 1028 1043 BY SIMILARITY.
FT DISULFID 1045 1054 BY SIMILARITY.
FT DISULFID 1061 1072 BY SIMILARITY.
FT DISULFID 1066 1081 BY SIMILARITY.
FT DISULFID 1083 1092 BY SIMILARITY.
FT DISULFID 1099 1120 BY SIMILARITY.

Query Match
Best Local Similarity 4.1%; Score 9; DB 1; Length 2437;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 NPGCHGTC 89
    |||||
Db 683 NPGCHGTC 691

RESULT 5
NTCL_MOUSE STANDARD; PRT; 2531 AA.
AC Q01705;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DE 01-OCT-2000 (Rel. 40, Last annotation update)
DE NEUROGENIC LOCUS NOTCH HOMOLOG PROTEIN 1 PRECURSOR (NOTCH PROTEIN).
GN NOTCH1 OR NOTCH.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=93194170; PubMed=8449489;
RA Franco del Amo F., Gendron-Maguire M., Swiatek P.J., Jenkins N.A.,
RA Copeland N.G., Gridley T.;
RT "Cloning, analysis, and chromosomal localization of Notch-1, a mouse
RL homolog of Drosophila Notch.";
RL Genomics 15:259-264(1993).
RN [2]
RP SEQUENCE OF 1551-2170 FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=93048835; PubMed=1425352;
RA Franco del Amo F., Smith D.E., Swiatek P.J., Gendron-Maguire M.,
RA Greenspan R.J., McMahon A.P., Gridley T.;
RT "Expression pattern of Notch, a mouse homolog of Drosophila Notch,
RT suggests an important role in early postimplantation mouse
RL development.";
RL Development 115:737-744(1992).
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- DEVELOPMENTAL STAGE: EXPRESSED ALMOST UNIFORMLY IN EARLY EMBRYOS.
CC -!- SIMILARITY: CONTAINS 36 EGF-LIKE DOMAINS.
CC -!- SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
CC -!- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC -!- SIMILARITY: HIGH, WITH OTHER NOTCH-TYPE PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch.
CC -----
CC EMBL; Z11886; CAA77941.1; -
CC HSP; P00740; 11XA.
DR MGD; MGI:97363; Notch1.
DR InterPro; IPR000152; -.
DR InterPro; IPR000361; -.
DR InterPro; IPR000800; -.
DR InterPro; IPR001438; -.
DR InterPro; IPR001881; -.
DR InterPro; IPR002110; -.
DR Pfam; PF00008; EGF; 35.
DR Pfam; PF00023; ank; 6.
DR Pfam; PF00066; notch; 3.
DR PRINTS; PR00010; EGFBL00D.
DR PROSITE; PS00088; ANK_REPEAT; 2.
DR PROSITE; PS00297; ANK_REPEAT_REGION; 1.
DR PROSITE; PS00010; ASX_HYDROXYL; 22.
DR PROSITE; PS00022; EGF_1; 34.
DR PROSITE; PS01186; EGF_2; 27.
DR PROSITE; PS01187; EGF_CA; 21.
KW Differentiation; Neurogenesis; Repeat; ANK repeat; EGF-like domain;
KW Transmembrane; Signal; Glycoprotein.
FT SIGNAL 1 18 POTENTIAL.
FT CHAIN 19 2531 NEUROGENIC LOCUS NOTCH HOMOLOG PROTEIN 1.
FT DOMAIN 19 1725 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1726 1746 POTENTIAL.
FT DOMAIN 1747 2531 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 20 58 EGF-LIKE 1.
FT DOMAIN 59 99 EGF-LIKE 2.
FT DOMAIN 102 139 EGF-LIKE 3.
FT DOMAIN 140 176 EGF-LIKE 4.
FT DOMAIN 178 216 EGF-LIKE 5.
FT DOMAIN 218 255 EGF-LIKE 6.
FT DOMAIN 257 293 EGF-LIKE 7.
FT DOMAIN 295 333 EGF-LIKE 8.
FT DOMAIN 335 371 EGF-LIKE 9.
FT DOMAIN 372 410 EGF-LIKE 10.
FT DOMAIN 412 450 EGF-LIKE 11.
FT DOMAIN 452 488 EGF-LIKE 12.
FT DOMAIN 490 526 EGF-LIKE 13.
FT DOMAIN 528 564 EGF-LIKE 14.
FT DOMAIN 566 601 EGF-LIKE 15.
FT DOMAIN 603 639 EGF-LIKE 16.
FT DOMAIN 641 676 EGF-LIKE 17.
FT DOMAIN 678 714 EGF-LIKE 18.
FT DOMAIN 716 751 EGF-LIKE 19.
FT DOMAIN 753 789 EGF-LIKE 20.
FT DOMAIN 791 827 EGF-LIKE 21.
FT DOMAIN 829 867 EGF-LIKE 22.
FT DOMAIN 869 905 EGF-LIKE 23.
FT DOMAIN 907 943 EGF-LIKE 24.
FT DOMAIN 945 981 EGF-LIKE 25.
FT DOMAIN 983 1019 EGF-LIKE 26.
FT DOMAIN 1021 1057 EGF-LIKE 27.
FT DOMAIN 1059 1095 EGF-LIKE 28.
FT DOMAIN 1097 1143 EGF-LIKE 29.
FT DOMAIN 1145 1181 EGF-LIKE 30.
FT DOMAIN 1183 1219 EGF-LIKE 31.
FT DOMAIN 1221 1265 EGF-LIKE 32.
FT DOMAIN 1267 1305 EGF-LIKE 33.
FT DOMAIN 1307 1346 EGF-LIKE 34.
FT DOMAIN 1348 1384 EGF-LIKE 35.
FT DOMAIN 1387 1426 EGF-LIKE 36.
FT DOMAIN 1449 1462 CYS-RICH.
FT REPEAT 1445 1480 LIN/NOTCH 1.
FT REPEAT 1481 1522 LIN/NOTCH 2.
FT REPEAT 1523 1562 LIN/NOTCH 3.
FT REPEAT 1917 1947 ANK 1.
FT REPEAT 1949 1979 ANK 2.
FT REPEAT 1983 2012 ANK 3.
FT REPEAT 2016 2045 ANK 4.
FT REPEAT 2049 2078 ANK 5.
FT DISULFID 24 37 BY SIMILARITY.
FT DISULFID 31 46 BY SIMILARITY.
FT DISULFID 63 74 BY SIMILARITY.
FT DISULFID 68 87 BY SIMILARITY.
FT DISULFID 89 98 BY SIMILARITY.
FT DISULFID 106 117 BY SIMILARITY.
FT DISULFID 111 127 BY SIMILARITY.
FT DISULFID 129 138 BY SIMILARITY.
FT DISULFID 144 155 BY SIMILARITY.

```

FT	DISULFID	1133	1142	BY SIMILARITY.
FT	DISULFID	1149	1160	BY SIMILARITY.
Query Match 4.1%; Score 9; DB 1; Length 2531;				
Best Local Similarity 100.0%; Pred. No. 0.56;				
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
QY	82	PCHNGGTCE	90	
Db	686	PCHNGGTCE	694	
RESULT 6				
ID	NTCL	RAT	STANDARD;	PRT; 2531 AA.
AC	Q07008;			
DT	01-NOV-1995	(Rel. 32, Created)		
DT	15-JUL-1999	(Rel. 38, Last sequence update)		
DT	01-OCT-2000	(Rel. 40, Last annotation update)		
DE	NEUROGENIC LOCUS NOTCH HOMOLOG PROTEIN 1 PRECURSOR.			
GN	NOTCH1.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.			
OX	NCBI_TaxID=10116;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE-Schwann cell;			
RX	MEDLINE=92111383; PubMed=1764995;			
RA	Weinmaster G., Roberts V.J., Lemke G.;			
RT	"A homolog of Drosophila Notch expressed during mammalian			
RT	development."			
RL	Development 113:199-205(1991).			
RN	[2]			
RP	REVISIONS TO 1652-1653.			
RA	Weinmaster G.;			
RL	Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.			
CC	-!- FUNCTION: REQUIRED FOR THE CORRECT DIFFERENTIATION OF A NUMBER			
CC	OF TISSUES.			
CC	-!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.			
CC	-!- DEVELOPMENTAL STAGE: IN THE EMBRYO, HIGHEST LEVELS OCCUR BETWEEN			
CC	DAYS 12 AND 14 AND DECREASE RAPIDLY TO MUCH LOWER LEVELS IN THE			
CC	ADULT.			
CC	-!- SIMILARITY: HIGH, WITH OTHER NOTCH-TYPE PROTEINS.			
CC	-!- SIMILARITY: CONTAINS 36 EGF-LIKE DOMAINS.			
CC	-!- SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.			
CC	-!- SIMILARITY: CONTAINS 5 ANK REPEATS.			

CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			

DR	EMBL; X57405; CAA40667.1;			
DR	HSSP; P00740; 11XA.			
DR	InterPro: IPR000152;			
DR	InterPro: IPR000561;			
DR	InterPro: IPR000800;			
DR	InterPro: IPR001438;			
DR	InterPro: IPR001881;			
DR	InterPro: IPR002049;			
DR	InterPro: IPR002110;			
DR	Pfam; PF00008; EGF; 36.			
DR	Pfam; PF00023; ank; 6.			
DR	Pfam; PF00066; notch; 3.			
DR	PRINTS; PR00010; EGFBL00D.			
DR	PRINTS; PR00011; EGFLAMININ.			
DR	PROSITE; PS50088; ANK_REPEAT; 4.			
DR	PROSITE; PS50297; ANK_REPEAT_REGION; 1.			
DR	PROSITE; PS00010; ASX_HYDROXYL; 22.			

DR	PROSITE; PS00022; EGF_1; 35.	FT	DISULFID	299	312	BY SIMILARITY.
DR	PROSITE; PS01186; EGF_2; 26.	FT	DISULFID	306	321	BY SIMILARITY.
DR	PROSITE; PS01187; EGF_CA; 21.	FT	DISULFID	323	332	BY SIMILARITY.
KW	Differentiation; Neurogenesis; Repeat; ANK repeat; EGF-like domain;	FT	DISULFID	339	350	BY SIMILARITY.
KW	Transmembrane; Signal; Glycoprotein.	FT	DISULFID	344	359	BY SIMILARITY.
	SIGNAL 1 18	FT	DISULFID	361	370	BY SIMILARITY.
	18	FT	DISULFID	376	387	BY SIMILARITY.
	CHAIN	FT	DISULFID	381	398	BY SIMILARITY.
	19 2531	FT	DISULFID	400	409	BY SIMILARITY.
	19 1723	FT	DISULFID	416	429	BY SIMILARITY.
	1724 1746	FT	DISULFID	423	438	BY SIMILARITY.
	TRANSMEM	FT	DISULFID	440	449	BY SIMILARITY.
	1747 2531	FT	DISULFID	456	467	BY SIMILARITY.
	DOMAIN	FT	DISULFID	461	476	BY SIMILARITY.
	20 58	FT	DISULFID	478	487	BY SIMILARITY.
	59 99	FT	DISULFID	494	505	BY SIMILARITY.
	DOMAIN	FT	DISULFID	499	514	BY SIMILARITY.
	102 139	FT	DISULFID	516	525	BY SIMILARITY.
	DOMAIN	FT	DISULFID	532	543	BY SIMILARITY.
	178 216	FT	DISULFID	537	552	BY SIMILARITY.
	218 255	FT	DISULFID	554	563	BY SIMILARITY.
	DOMAIN	FT	DISULFID	570	580	BY SIMILARITY.
	257 293	FT	DISULFID	575	589	BY SIMILARITY.
	DOMAIN	FT	DISULFID	591	600	BY SIMILARITY.
	295 333	FT	DISULFID	607	618	BY SIMILARITY.
	DOMAIN	FT	DISULFID	629	638	BY SIMILARITY.
	335 371	FT	DISULFID	645	655	BY SIMILARITY.
	372 410	FT	DISULFID	666	675	BY SIMILARITY.
	412 450	FT	DISULFID	682	693	BY SIMILARITY.
	452 488	FT	DISULFID	687	702	BY SIMILARITY.
	DOMAIN	FT	DISULFID	704	713	BY SIMILARITY.
	490 526	FT	DISULFID	720	730	BY SIMILARITY.
	DOMAIN	FT	DISULFID	725	739	BY SIMILARITY.
	566 601	FT	DISULFID	741	750	BY SIMILARITY.
	DOMAIN	FT	DISULFID	757	768	BY SIMILARITY.
	603 639	FT	DISULFID	762	777	BY SIMILARITY.
	641 676	FT	DISULFID	779	788	BY SIMILARITY.
	DOMAIN	FT	DISULFID	795	806	BY SIMILARITY.
	678 714	FT	DISULFID	800	815	BY SIMILARITY.
	716 751	FT	DISULFID	817	826	BY SIMILARITY.
	DOMAIN	FT	DISULFID	833	844	BY SIMILARITY.
	753 789	FT	DISULFID	838	855	BY SIMILARITY.
	791 827	FT	DISULFID	857	866	BY SIMILARITY.
	829 867	FT	DISULFID	873	884	BY SIMILARITY.
	EGF-LIKE 21.	FT	DISULFID	878	893	BY SIMILARITY.
	EGF-LIKE 22.	FT	DISULFID	895	904	BY SIMILARITY.
	EGF-LIKE 23.	FT	DISULFID	911	922	BY SIMILARITY.
	EGF-LIKE 24.	FT	DISULFID	916	931	BY SIMILARITY.
	EGF-LIKE 25.	FT	DISULFID	933	942	BY SIMILARITY.
	EGF-LIKE 26.	FT	DISULFID	942	951	BY SIMILARITY.
	EGF-LIKE 27.	FT	DISULFID	951	960	BY SIMILARITY.
	EGF-LIKE 28.	FT	DISULFID	960	969	BY SIMILARITY.
	EGF-LIKE 29.	FT	DISULFID	969	978	BY SIMILARITY.
	EGF-LIKE 30.	FT	DISULFID	978	987	BY SIMILARITY.
	EGF-LIKE 31.	FT	DISULFID	987	996	BY SIMILARITY.
	EGF-LIKE 32.	FT	DISULFID	996	1005	BY SIMILARITY.
	EGF-LIKE 33.	FT	DISULFID	1005	1014	BY SIMILARITY.
	EGF-LIKE 34.	FT	DISULFID	1014	1023	BY SIMILARITY.
	EGF-LIKE 35.	FT	DISULFID	1023	1032	BY SIMILARITY.
	EGF-LIKE 36.	FT	DISULFID	1032	1041	BY SIMILARITY.
	CYS-RICH.	FT	DISULFID	1041	1050	BY SIMILARITY.
	ANK 1.	FT	DISULFID	1050	1059	BY SIMILARITY.
	1917 1946	FT	DISULFID	1059	1068	BY SIMILARITY.
	REPEAT	FT	DISULFID	1068	1077	BY SIMILARITY.
	1950 1980	FT	DISULFID	1077	1086	BY SIMILARITY.

Qy 82 PCHNGTCE 90
 Db 686 PCHNGTCE 694

RESULT 7
 DL_DROME STANDARD; PRT; 833 AA.
 ID DL_DROME
 AC P10041; Q9VDY2; 099108;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE NEUROGENIC LOCUS DELTA PROTEIN PRECURSOR.
 GN DL OR CG3619.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A., TISSUE SPECIFICITY, AND DEVELOPMENTAL STAGE.
 RC TISSUE=Embryo;
 RA Vaessin H., Bremer K.A., Knust E., Campos-Ortega J.A.;
 RT "The neurogenic gene Delta of Drosophila melanogaster is expressed in
 RT neurogenic territories and encodes a putative transmembrane protein
 RT with EGF-like repeats.";
 RL EMBO J. 6:3431-3440(1987).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=OREGON-R; TISSUE=Embryo;
 RX MEDLINE=89196890; PubMed=3149249;
 RA Kocyanski C.C., Alton A.K., Fechtel K., Koeh P.J., Muskavitch M.A.T.;
 RT "Delta, a Drosophila neurogenic gene, is transcriptionally complex and
 RT encodes a protein related to blood coagulation factors and epidermal
 RT growth factor of vertebrates.";
 RL Genes Dev. 2:1723-1735(1988).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Ananthides P.G., Scher S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Vandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers J., H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
 RA Abril J.F., Adayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K.J., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
 RA Flossler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,

RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 RN [4]
 RP SEQUENCE OF 422-621 FROM N.A.
 RC TISSUE=Embryo;
 RX MEDLINE=87218537; PubMed=3107986;
 RA Knust E., Dietrich U., Tepass U., Bremer K.A., Weigel D., Vaessin H.,
 RA Campos-Ortega J.A.;
 RT "EGF homologous sequences encoded in the genome of Drosophila
 RT melanogaster, and their relation to neurogenic genes.";
 RL EMBO J. 6:761-766(1987).
 RN [5]
 RP PATTERN OF TRANSCRIPTION, AND CHARACTERIZATION.
 RX MEDLINE=91209246; PubMed=2128477;
 RA Haenlin M., Kramatschek B., Campos-Ortega J.A.;
 RT "The pattern of transcription of the neurogenic gene Delta of
 RT Drosophila melanogaster.";
 RL Development 110:905-914(1990).
 CC -!- FUNCTION: ESSENTIAL FOR PROPER DIFFERENTIATION OF ECTODERM. DL
 CC IS REQUIRED FOR THE CORRECT SEPARATION OF NEURAL AND EPIDERMAL
 CC CELL LINEAGES.
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -!- TISSUE SPECIFICITY: DETECTED IN ALL AREAS WITH NEUROGENIC
 CC ABILITIES, FOR EXAMPLE THE NEUROGENIC ECTODERM AND THE PRIMORDIA
 CC OF THE SENSE ORGANS. LATER EXPRESSION IS RESTRICTED TO THOSE CELLS
 CC THAT HAVE ADOPTED A NEURAL FATE.
 CC -!- DEVELOPMENTAL STAGE: EXPRESSED BOTH MATERNALLY AND ZYGOTICALLY.
 CC REDUCES TO A LOW LEVEL EARLY IN EMBRYONIC DEVELOPMENT (STAGE 5) AND
 CC EXPRESSION IS HIGHEST DURING LARVAL STAGES.
 CC -!- MISCELLANEOUS: SEPARATION OF NEUROBLASTS FROM THE ECTODERM INTO
 CC THE INNER PART OF EMBRYO IS ONE OF THE FIRST STEPS OF CNS
 CC DEVELOPMENT IN INSECTS, THIS PROCESS IS UNDER CONTROL OF THE
 CC NEUROGENIC GENES.
 CC -!- MISCELLANEOUS: NOTCH AND SERRATE MAY INTERACT AT THE PROTEIN
 CC LEVEL, IT IS CONCEIVABLE THAT THE SERRATE AND DELTA PROTEINS MAY
 CC COMPETE FOR BINDING WITH THE NOTCH PROTEIN.
 CC -!- SIMILARITY: CONTAINS 9 EGF-LIKE DOMAINS.
 CC -!- SIMILARITY: BELONGS TO THE DELTA/SERRATE/JAGGED FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X06289; CAA29617.1; -
 DR EMBL; Y00222; CAA68369.1; -
 DR EMBL; AE003725; AAF55657.1; -
 DR EMBL; X05140; CAA28786.1; -
 DR PIR; S00670; S00670.
 DR PIR; A26637; A26637.
 DR HSP; P00740; IIXA.
 DR FlyBase; FBgn0000463; D1.
 DR InterPro; IPR000152; -
 DR InterPro; IPR000561; -
 DR InterPro; IPR001774; -
 DR InterPro; IPR001881; -
 DR Pfam; PF01414; DSL; 1.
 DR Pfam; PF00008; EGF; 9.
 DR PROSITE; PS00010; ASX_HYDROXYL; 3.
 DR PROSITE; PS00022; EGF_1; 9.
 DR PROSITE; PS01186; EGF_2; 9.
 DR PROSITE; PS01187; EGF_CA; 2.
 KW Differentiation; Neurogenesis; Repeat; Transmembrane;
 KW EGF-like domain; Glycoprotein; Signal.
 FT SIGNAL 1 18 POTENTIAL.
 FT CHAIN 19 833 NEUROGENIC LOCUS DELTA PROTEIN.
 FT DOMAIN 19 653 EXTRACELLULAR (POTENTIAL).


```

FT DISULFID 202 223 BY SIMILARITY.
FT DISULFID 276 351 BY SIMILARITY.
FT DISULFID 300 321 BY SIMILARITY.
FT DISULFID 626 637 BY SIMILARITY.
FT DISULFID 631 646 BY SIMILARITY.
FT DISULFID 648 657 BY SIMILARITY.
FT DISULFID 664 675 BY SIMILARITY.
FT DISULFID 692 784 BY SIMILARITY.
FT DISULFID 760 776 BY SIMILARITY.
FT DISULFID 791 834 BY SIMILARITY.
FT DISULFID 820 847 BY SIMILARITY.
FT CARBOHYD 129 129 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 336 336 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 883 AA; 96013 MW; CC2C33C97B453E45 CRC64;

Query Match 3 6%; Score 8; DB 1; Length 883;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 82 PCHNGGTC 89
| | | | | | | |
DB 630 PCHNGGTC 637

RESULT 9
PCGB_RAT
ID PCGB_RAT STANDARD; PRT: 883 AA.
AC P5068; Q63040; Q62860; Q63513;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE BREVICAN CORE PROTEIN PRECURSOR (BRAIN ENRICHED HYALURONAN BINDING
DE PROTEIN) (BEHAV PROTEIN).
GN BCAN.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Brain;
RX MEDLINE=96070828; PubMed=7592978;
RA Seidenbecher C.I., Richter K., Rauch U., Faessler R., Garner C.C.,
RA Gundelfinger E.D.;
RT "Brevican, a chondroitin sulfate proteoglycan of rat brain, occurs as
RT secreted and cell surface glycosylphosphatidylinositol-anchored
RT isoforms."
RL J. Biol. Chem. 270:27206-27212(1995).
RN [2]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 396-407.
RC TISSUE=Brain;
RX MEDLINE=96074575; PubMed=7488217;
RA Yamada H., Watanabe K., Shimonaka M., Yamasaki M., Yamauchi Y.;
RT "cDNA cloning and the identification of an aggrecanase-like cleavage
RT site in rat brevicane."
RL Biochem. Biophys. Res. Commun. 216:957-963(1995).
RN [3]
RP SEQUENCE OF 1-423 FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Brain;
RX MEDLINE=94216386; PubMed=7512973;
RA Jaworski D.M., Kelly G.M., Hockfield S.;
RT "BEHAV, a new member of the proteoglycan tandem repeat family of
RT hyaluronan-binding proteins that is restricted to the brain."
RL J. Cell Biol. 125:495-509(1994).
CC -1- FUNCTION: MAY PLAY A ROLE IN THE TERMINALLY DIFFERENTIATING AND
CC THE ADULT NERVOUS SYSTEM DURING POSTNATAL DEVELOPMENT. COULD
CC STABILIZE INTERACTIONS BETWEEN HA AND BRAIN PROTEOGLYCANS. THE
CC GPI-ANCHORED ISOFORM MAY FUNCTION AS A CHONDROITIN SULFATE-
CC BEARING CELL SURFACE RECEPTOR.
CC -1- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR MATRIX AND ONE FORM
CC ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR.
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A SECRETED FORM (SHOWN HERE) AND

```

```

CC A GPI-ANCHORED FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: BRAIN.
CC -1- DEVELOPMENTAL STAGE: SOLUBLE FORM INCREASES FROM DAY P4 TO P64.
CC GPI-ANCHORED ISOFORM INCREASES AFTER DAY P8.
CC -1- PTM: CONTAINS MOSTLY CHONDROITIN SULFATE.
CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.
CC -1- SIMILARITY: CONTAINS 2 LINK DOMAINS.
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 SUSHI (SCR) REPEAT.
CC -1- SIMILARITY: BELONGS TO THE AGGRECAN/VERSICAN PROTEOGLYCAN FAMILY.
CC -1- CAUTION: REF.3 SEQUENCE DIFFERS FROM THAT SHOWN FROM POSITION 364
CC ONWARD AND IS SMALLER (371 AA) DUE TO A FRAMESHIFT.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X79881; CAA56255.1; -
CC EMBL; X86406; CAA60160.1; -
CC EMBL; U37142; CAA87847.1; -
CC EMBL; Z28366; CAA82215.1; ALT_FRAME.
CC HSSP; P20693; 1HLJ.
CC InterPro; IPR000436; -
CC InterPro; IPR000495; -
CC InterPro; IPR000538; -
CC InterPro; IPR000561; -
CC InterPro; IPR001304; -
CC InterPro; IPR003006; -
CC Pfam; PF00008; EGF_1; 1.
CC Pfam; PF00193; Xlink; 2.
CC Pfam; PF00047; ig; 1.
CC Pfam; PF00059; lectin_c; 1.
CC Pfam; PF00084; sushi; 1.
CC PROSITE; PS00022; EGF_1; 1.
CC PROSITE; PS01186; EGF_2; 1.
CC PROSITE; PS00290; IG_MHC; 1.
CC PROSITE; PS01241; LINK; 2.
CC PROSITE; PS00615; C-TYPE_LECTIN_1; 1.
CC PROSITE; PS00041; C-TYPE_LECTIN_2; 1.
CC Glycoprotein; Hyaluronic acid; Proteoglycan; Lectin; Signal; Sushi;
CC EGF-like domain; Repeat; Immunoglobulin domain; Alternative splicing;
CC GPI-anchor.
CC SIGNAL 1 22 POTENTIAL.
CC CHAIN 23 883 BREVICAN CORE PROTEIN.
CC DOMAIN 32 157 IG-LIKE V-TYPE DOMAIN.
CC DOMAIN 173 250 LINK 1.
CC DOMAIN 271 352 LINK 2.
CC DOMAIN 622 658 EGF-LIKE.
CC DOMAIN 658 786 C-TYPE LECTIN.
CC DOMAIN 787 851 SUSHI.
CC DISULFID 56 136 BY SIMILARITY.
CC DISULFID 178 249 BY SIMILARITY.
CC DISULFID 202 223 BY SIMILARITY.
CC DISULFID 276 351 BY SIMILARITY.
CC DISULFID 300 321 BY SIMILARITY.
CC DISULFID 626 637 BY SIMILARITY.
CC DISULFID 648 657 BY SIMILARITY.
CC DISULFID 791 834 BY SIMILARITY.
CC DISULFID 820 847 BY SIMILARITY.
CC CARBOHYD 129 129 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 336 336 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC VARSPPLIC 625 645 DCIPSPCHNGGTCLEKEGFR -> NSAEKGNPAFLFLLL
CC VARSPPLIC 646 883 QMWDT (IN GPI-ANCHORED ISOFORM).
CC VARSPPLIC 51 52 MISSING (IN GPI-ANCHORED ISOFORM).
CC CONFLICT 503 503 AL -> WV (IN REF. 2).
CC CONFLICT 518 519 V -> L (IN REF. 3).
CC CONFLICT 519 519 TV -> PA (IN REF. 2).

```


FT CONFLICT 526 526 G -> R (IN REF. 2).
FT CONFLICT 541 541 G -> A (IN REF. 2).
FT CONFLICT 556 556 R -> S (IN REF. 2).
FT CONFLICT 573 573 E -> A (IN REF. 2).
FT CONFLICT 583 583 V -> L (IN REF. 2).
FT CONFLICT 649 649 V -> L (IN REF. 2).
FT CONFLICT 670 670 P -> A (IN REF. 2).
FT CONFLICT 738 738 P -> A (IN REF. 2).
FT CONFLICT 809 809 R -> A (IN REF. 2).
SQ SEQUENCE 883 AA; 96057 MW; AC/ACC40CB53ED37 CRC64;

Query Match 3.6%; Score 8; DB 1; Length 883;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 82 PCHNGGTC 89
|||||||
Db 630 PCHNGGTC 637

RESULT 10
PCGB_BOVIN STANDARD; PRT; 912 AA.
AC Q28062;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE BREVICAN CORE PROTEIN PRECURSOR.
GN BCAN.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=94193597; PubMed=8144512;
RA Yamada H., Watanabe K., Shimonaka M., Yamaguchi Y.;
RT "Molecular cloning of Brevican, a novel brain proteoglycan of the aggrecan/versican family.";
RL J. Biol. Chem. 269:10119-10126(1994).
CC -1- FUNCTION: MAY PLAY A ROLE IN THE TERMINALLY DIFFERENTIATING AND STABILIZE NERVOUS SYSTEM DURING POSTNATAL DEVELOPMENT. COULD STABILIZE INTERACTIONS BETWEEN HA AND BRAIN PROTEOGLYCANS.
CC -1- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR MATRIX (BY SIMILARITY).
CC -1- TISSUE SPECIFICITY: BRAIN; EXPRESSED IN CEREBELLAR ASTROCYTES BUT NOT IN NEURONS.
CC -1- PTM: CONTAINS MOSTLY CHONDROITIN SULFATE.
CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.
CC -1- SIMILARITY: CONTAINS 2 LINK DOMAINS.
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 SUSHI (SCR) REPEAT.
CC -1- SIMILARITY: BELONGS TO THE AGGREGAN/VERSICAN PROTEOGLYCAN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X75987; CAA53481.1; -;
CC HSPSP; P20693; IHLJ.
CC InterPro; IPR000436; -;
CC InterPro; IPR000495; -;
CC InterPro; IPR000538; -;
CC InterPro; IPR000561; -;
CC InterPro; IPR001304; -;

DR InterPro; IPR003006; -;
DR Pfam; PF00008; EGF_1;
DR Pfam; PF00193; Xlink; 2.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF00059; lectin_c; 1.
DR Pfam; PF00084; sush1; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00290; IG_MHC; 1.
DR PROSITE; PS01241; LINK; 2.
DR PROSITE; PS00615; C-TYPE_LECTIN_1; 1.
DR PROSITE; PS00041; C-TYPE_LECTIN_2; 1.
KW Glycoprotein; Hyaluronic acid; Proteoglycan; Lectin; Signal; Sushi;
KW EGF-like domain; Repeat; Immunoglobulin domain.
FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 912 BREVICAN CORE PROTEIN.
FT DOMAIN 32 158 IG-LIKE V-TYPE DOMAIN.
FT DOMAIN 174 251 LINK 1.
FT DOMAIN 272 353 LINK 2.
FT DOMAIN 647 683 EGF-LIKE.
FT DOMAIN 683 811 C-TYPE LECTIN.
FT DOMAIN 812 876 SUSHI.
FT DISULFID 57 137 BY SIMILARITY.
FT DISULFID 179 250 BY SIMILARITY.
FT DISULFID 203 224 BY SIMILARITY.
FT DISULFID 277 352 BY SIMILARITY.
FT DISULFID 301 322 BY SIMILARITY.
FT DISULFID 651 662 BY SIMILARITY.
FT DISULFID 656 671 BY SIMILARITY.
FT DISULFID 673 682 BY SIMILARITY.
FT DISULFID 689 700 BY SIMILARITY.
FT DISULFID 717 809 BY SIMILARITY.
FT DISULFID 785 801 BY SIMILARITY.
FT DISULFID 816 859 BY SIMILARITY.
FT DISULFID 845 872 BY SIMILARITY.
FT CARBOHYD 130 130 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 337 337 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 912 AA; 99554 MW; 677B3EB1C688C4D7 CRC64;

Query Match 3.6%; Score 8; DB 1; Length 912;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 82 PCHNGGTC 89
|||||||
Db 655 PCHNGGTC 662

RESULT 11
NOTC_XENLA STANDARD; PRT; 2524 AA.
ID NOTC_XENLA
AC P21783;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE NEUROGENIC LOCUS NOTCH PROTEIN HOMOLOG PRECURSOR (XOTCH PROTEIN).
GN XOTCH.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90385285; PubMed=2402639;
RA Coffman C., Harris W., Kintner C.;
RT "Xotch, the Xenopus homolog of Drosophila notch.";
RL Science 249:1438-1441(1990).
RN [2]
RP REVISIONS TO 1759-1782.
RA Kintner C.;
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.

FT DISULFID 794 805 BY SIMILARITY.
FT DISULFID 799 814 BY SIMILARITY.
FT DISULFID 816 825 BY SIMILARITY.
FT DISULFID 832 843 BY SIMILARITY.
FT DISULFID 837 854 BY SIMILARITY.
FT DISULFID 856 865 BY SIMILARITY.
FT DISULFID 872 883 BY SIMILARITY.
FT DISULFID 877 892 BY SIMILARITY.
FT DISULFID 894 903 BY SIMILARITY.
FT DISULFID 910 921 BY SIMILARITY.
FT DISULFID 915 930 BY SIMILARITY.
FT DISULFID 932 941 BY SIMILARITY.
FT DISULFID 986 997 BY SIMILARITY.
FT DISULFID 991 1006 BY SIMILARITY.
FT DISULFID 1008 1017 BY SIMILARITY.
FT DISULFID 1024 1035 BY SIMILARITY.
FT DISULFID 1029 1044 BY SIMILARITY.
FT DISULFID 1046 1055 BY SIMILARITY.
FT DISULFID 1062 1073 BY SIMILARITY.
FT DISULFID 1067 1082 BY SIMILARITY.
FT DISULFID 1084 1093 BY SIMILARITY.
FT DISULFID 1100 1121 BY SIMILARITY.
FT DISULFID 1115 1130 BY SIMILARITY.
FT DISULFID 1132 1141 BY SIMILARITY.
FT DISULFID 1148 1159 BY SIMILARITY.
FT DISULFID 1153 1168 BY SIMILARITY.
FT DISULFID 1170 1179 BY SIMILARITY.
FT DISULFID 1186 1197 BY SIMILARITY.
FT DISULFID 1191 1206 BY SIMILARITY.
FT DISULFID 1208 1217 BY SIMILARITY.
FT DISULFID 1224 1243 BY SIMILARITY.
FT DISULFID 1237 1252 BY SIMILARITY.

Query Match 3.6%; Score 8; DB 1; Length 2524;
Best Local Similarity 100.0%; Pred. No. 6.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 PNPCHNGG 87
Db 912 PNPCHNGG 919

RESULT 12
YAQ5_SCHPO
ID YAQ5_SCHPO STANDARD; PRT; 2670 AA.
AC Q10105;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE PUTATIVE TRANSLATIONAL ACTIVATOR C18G6.05C (GCN1 HOMOLOG).
GN SPAC18G6.05C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomyces.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Connor R., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: STRONG, TO YEAST GCN1.
CC -1- SIMILARITY: CONTAINS 19 HEAT REPEATS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z68198; CAA92385.1; -

DR PROSITE; PS50077; HEAT_REPEAT; 4.
KW Hypothetical protein; Translation regulation; Activator; Repeat.
FT REPEAT 315 352 HEAT 1.
FT REPEAT 1062 1099 HEAT 2.
FT REPEAT 1319 1356 HEAT 3.
FT REPEAT 1439 1476 HEAT 4.
FT REPEAT 1478 1514 HEAT 5.
FT REPEAT 1518 1555 HEAT 6.
FT REPEAT 1557 1593 HEAT 7.
FT REPEAT 1637 1674 HEAT 8.
FT REPEAT 1676 1713 HEAT 9.
FT REPEAT 1714 1751 HEAT 10.
FT REPEAT 1755 1792 HEAT 11.
FT REPEAT 1793 1830 HEAT 12.
FT REPEAT 1898 1939 HEAT 13.
FT REPEAT 1941 1977 HEAT 14.
FT REPEAT 1982 2019 HEAT 15.
FT REPEAT 2020 2055 HEAT 16.
FT REPEAT 2057 2090 HEAT 17.
FT REPEAT 2319 2361 HEAT 18.
FT REPEAT 2379 2416 HEAT 19.
SQ SEQUENCE 2670 AA; 297333 MW; 948E9316D56D74C3 CRC64;

Query Match 3.6%; Score 8; DB 1; Length 2670;
Best Local Similarity 100.0%; Pred. No. 6.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 38 LSLGADDS 45
Db 1765 LSLGADDS 1772

RESULT 13
SH_RAT
ID SH_RAT STANDARD; PRT; 106 AA.
AC P55248;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE PUTATIVE PROTEIN SH (SH-4).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Heart;
RX MEDLINE=87149087; PubMed=3547652;
RA Adelman J.P., Bond C.T., Douglass J., Herbert E.;
RT "Two mammalian genes transcribed from opposite strands of the same
RT DNA locus".
RL Science 235:1514-1517(1987).
CC -1- FUNCTION: MAY BE INVOLVED WITH THE REGULATION OF GNRH GENE
CC EXPRESSION. IT IS NOT KNOWN IF THIS PROTEIN IS TRANSCRIBED.
CC -1- TISSUE SPECIFICITY: HEART.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M15527; AAA42140.1; -
SQ SEQUENCE 106 AA; 11792 MW; F9EF8FDBFEAABE3E CRC64;

Query Match 3.2%; Score 7; DB 1; Length 106;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 33 NGGICLS 39
Db 100 NGGICLS 106

RESULT 14
CYOD_PSEPU
ID CYOD_PSEPU STANDARD; PRT; 110 AA.
AC 09WWR4;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE CYTOCHROME O UBIQUINOL OXIDASE PROTEIN CYOD.
GN CYOD.
OS Pseudomonas putida.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IH-2000;
RX MEDLINE=99085656; PubMed=9868765;
RA Hirayama H., Takami H., Inoue A., Horikoshi K.;
RT "Isolation and characterization of toluene-sensitive mutants from
RT Pseudomonas putida IH-2000.";
RL FEMS Microbiol. Lett. 169:219-225(1998).
CC -!- FUNCTION: CYTOCHROME O TERMINAL OXIDASE COMPLEX IS THE COMPONENT
CC OF THE AEROBIC RESPIRATORY CHAIN THAT PREDOMINATES WHEN CELLS ARE
CC GROWN AT HIGH AERATION (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE
CC (BY SIMILARITY).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB016787; BAA76359.1;
DR DOMAIN 1 18 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 19 39 POTENTIAL.
FT DOMAIN 40 46 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 47 67 POTENTIAL.
FT DOMAIN 68 78 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 79 99 POTENTIAL.
FT DOMAIN 100 110 PERIPLASMIC (POTENTIAL).
SQ SEQUENCE 110 AA; 12376 MW; 074ED0BE854FC0C9 CRC64;

Query Match 3.2%; Score 7; DB 1; Length 110;
Best Local Similarity 100.0%; Pred. No. 4.6;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 LVLGLSL 15
Db 91 LVLGLSL 97

RESULT 15
CRAA_RANTE STANDARD; PRT; 149 AA.
AC P02508;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE ALPHA CRYSTALLIN A CHAIN (FRAGMENT).
GN CRYAA.
OS Rana temporaria (European common frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranioidea; Ranidae; Rana.

Query Match 3.2%; Score 7; DB 1; Length 149;
Best Local Similarity 100.0%; Pred. No. 6;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 69 EERPTSA 75
Db 140 EERPTSA 146

RESULT 16
CRAA_RANES STANDARD; PRT; 167 AA.
AC P02507;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE ALPHA CRYSTALLIN A CHAIN (FRAGMENTS).
GN CRYAA.
OS Rana esculenta (Edible frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranioidea; Ranidae; Rana.
OX NCBI_TaxID=8401;
RN [1]
RP SEQUENCE.
RX MEDLINE=84208008; PubMed=6723655;
RA de Jong W.W., Zweers A., Versteeg M., Nuy-Terwindt E.C.;
RT "Primary structures of the alpha-crystallin A chains of twenty-eight
RT mammalian species, chicken and frog.";
RL Eur. J. Biochem. 141:131-140(1984).
CC -!- FUNCTION: MAY CONTRIBUTE TO THE TRANSPARENCY AND REFRACTIVE INDEX
CC OF THE LENS.
CC -!- SIMILARITY: BELONGS TO THE SMALL HEAT SHOCK PROTEIN (HSP20)
CC FAMILY. STRONG TO ALPHA(B)-CRYSTALLIN.
CC PIR: A02910; CYFGAA.
DR InterPro; IPR002068;
DR Pfam; PF00011; HSP20; 1.
DR Pfam; PF00525; crystallin; 1.
DR PROSITE; PS01031; HSP20; 1.
KW Eye lens protein; Acetylation; Glycoprotein.
FT MOD_RES 1 1 ACETYLATION.

NCBI_TaxID=8407;
OX [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=84005173; PubMed=6604666;
RA Tomarev S.I., Zinovleva R.D., Dolgilevich S.M., Kravayev A.S.,
RA Skryabin K.G., Gause G.G. Jr.;
RT "The absence of the long 3'-non-translated region in mRNA coding for
RT eye lens alpha A2-crystallin of the frog (Rana temporaria).";
RL FEBS Lett. 162:47-51(1983).
CC -!- FUNCTION: MAY CONTRIBUTE TO THE TRANSPARENCY AND REFRACTIVE INDEX
CC OF THE LENS.
CC -!- SIMILARITY: BELONGS TO THE SMALL HEAT SHOCK PROTEIN (HSP20)
CC FAMILY. STRONG TO ALPHA(B)-CRYSTALLIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X00716; CAA25308.1;
DR PIR; A02911; CYFGA2.
DR InterPro; IPR002068;
DR Pfam; PF00011; HSP20; 1.
DR PROSITE; PS01031; HSP20; 1.
KW Eye lens protein; Glycoprotein.
FT NON_TER 1 1
FT CARBOHYD 138 138 O-LINKED (GLCNAC) (BY SIMILARITY).
SQ SEQUENCE 149 AA; 16949 MW; 877E89A7428DBADC CRC64;

Query Match 3.2%; Score 7; DB 1; Length 149;
Best Local Similarity 100.0%; Pred. No. 6;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 69 EERPTSA 75
Db 140 EERPTSA 146

RESULT 16
CRAA_RANES STANDARD; PRT; 167 AA.
AC P02507;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE ALPHA CRYSTALLIN A CHAIN (FRAGMENTS).
GN CRYAA.
OS Rana esculenta (Edible frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranioidea; Ranidae; Rana.
OX NCBI_TaxID=8401;
RN [1]
RP SEQUENCE.
RX MEDLINE=84208008; PubMed=6723655;
RA de Jong W.W., Zweers A., Versteeg M., Nuy-Terwindt E.C.;
RT "Primary structures of the alpha-crystallin A chains of twenty-eight
RT mammalian species, chicken and frog.";
RL Eur. J. Biochem. 141:131-140(1984).
CC -!- FUNCTION: MAY CONTRIBUTE TO THE TRANSPARENCY AND REFRACTIVE INDEX
CC OF THE LENS.
CC -!- SIMILARITY: BELONGS TO THE SMALL HEAT SHOCK PROTEIN (HSP20)
CC FAMILY. STRONG TO ALPHA(B)-CRYSTALLIN.
CC PIR: A02910; CYFGAA.
DR InterPro; IPR002068;
DR Pfam; PF00011; HSP20; 1.
DR Pfam; PF00525; crystallin; 1.
DR PROSITE; PS01031; HSP20; 1.
KW Eye lens protein; Acetylation; Glycoprotein.
FT MOD_RES 1 1 ACETYLATION.

```

FT NON_CONS 70 71
 FT CARBOHYD 156 156 O-LINKED (GLCNAC) (BY SIMILARITY).
 SQ SEQUENCE 167 AA; 19242 MW; EEDCB4DD6EE5469A CRC64;

Query Match 3.2%; Score 7; DB 1; Length 167;
 Best Local Similarity 100.0%; Pred. No. 6.6;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 69 EEKPTSA 75
 |||||
 DB 158 EEKPTSA 164

RESULT 17
 CAA_AALMI STANDARD; PRT; 173 AA.
 AC P06904;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE ALPHA CRYSTALLIN A CHAIN.
 GN. CRYAA.
 OS Alligator mississippiensis (American alligator).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Crocodylidae; Alligatorinae; Alligator.
 OX NCBI_TaxID=8496;
 RN [1]
 RP PARTIAL SEQUENCE.

RX MEDLINE=88216135; PubMed=3870872;
 RA de Jong W.W., Zweers A., Versteeg M., Dessauer H.C., Goodman M.;
 RT "Alpha-crystallin A sequences of Alligator mississippiensis and the
 lizard Tupinambis teguixin: molecular evolution and reptilian
 phylogeny.";
 RL Mol. Biol. Evol. 2:484-493(1985).
 CC -!- FUNCTION: MAY CONTRIBUTE TO THE TRANSPARENCY AND REFRACTIVE INDEX
 OF THE LENS.
 CC -!- SIMILARITY: BELONGS TO THE SMALL HEAT SHOCK PROTEIN (HSP20)

CC FAMILY: STRONG TO ALPHA(B)-CRYSTALLIN.
 DR PIR; A25753; CYAQA.
 DR InterPro; IPR001436;
 DR InterPro; IPR002068;
 DR InterPro; IPR003090;
 DR Pfam; PF00011; HSP20; 1.
 DR PRINTS; PR00299; ACRYSTALLIN.
 DR PROSITE; PS01031; HSP20; 1.
 KW Eye lens protein; Acetylation.
 FT MOD_RES 1 1 ACETYLATION (PROBABLE).
 SQ SEQUENCE 173 AA; 19769 MW; 9B9DD0B549BD2C2A CRC64;

Query Match 3.2%; Score 7; DB 1; Length 173;
 Best Local Similarity 100.0%; Pred. No. 6.8;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 69 EEKPTSA 75
 |||||
 DB 164 EEKPTSA 170

RESULT 18
 CAA_CHICK STANDARD; PRT; 173 AA.
 AC P02504;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-JUL-1989 (Rel. 11, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE ALPHA CRYSTALLIN A CHAIN.
 GN. CRYAA.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;

OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88056321; PubMed=3678835;
 RA Thompson M.A., Hawkins J.W., Piatigorsky J.;
 RT "Complete nucleotide sequence of the chicken alpha A-crystallin gene
 and its 5' flanking region.";
 RL Gene 56:173-184(1987).
 RN [2]
 RP PRELIMINARY SEQUENCE.
 RX MEDLINE=84208008; PubMed=6723655;
 RA de Jong W.W., Zweers A., Versteeg M., Nuy-Terwindt E.C.;
 RT "Primary structures of the alpha-crystallin A chains of twenty-eight
 mammalian species, chicken and frog.";
 RL Eur. J. Biochem. 141:131-140(1984).
 CC -!- FUNCTION: MAY CONTRIBUTE TO THE TRANSPARENCY AND REFRACTIVE INDEX
 OF THE LENS.
 CC -!- SIMILARITY: BELONGS TO THE SMALL HEAT SHOCK PROTEIN (HSP20)

CC FAMILY: STRONG TO ALPHA(B)-CRYSTALLIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M17627; AAA48722.1;
 DR PIR; A02907; CYCHAA.
 DR PIR; A27309; A27309.
 DR InterPro; IPR001436;
 DR InterPro; IPR002068;
 DR InterPro; IPR003090;
 DR Pfam; PF00011; HSP20; 1.
 DR Pfam; PF00525; crystallin; 1.
 DR PRINTS; PR00299; ACRYSTALLIN.
 DR PROSITE; PS01031; HSP20; 1.
 KW Eye lens protein; Acetylation; Glycoprotein.
 FT MOD_RES 1 1 ACETYLATION (PROBABLE).
 FT CARBOHYD 162 162 O-LINKED (GLCNAC) (BY SIMILARITY).
 SQ SEQUENCE 173 AA; 19787 MW; 327AAFCAD379B399 CRC64;

Query Match 3.2%; Score 7; DB 1; Length 173;
 Best Local Similarity 100.0%; Pred. No. 6.8;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 69 EEKPTSA 75
 |||||
 DB 164 EEKPTSA 170

RESULT 19
 CAA_HUMAN STANDARD; PRT; 173 AA.
 AC P02489;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE ALPHA CRYSTALLIN A CHAIN.
 GN. CRYAA OR CRYAL.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lens;
 RX MEDLINE=96139023; PubMed=8587135;
 RA Jaworski C.J.;
 RT "A reassessment of mammalian alpha A-crystallin sequences using DNA

sequencing: implications for anthropoid affinities of tarsier.;"
J. Mol. Evol. 41:901-908(1995).
[2]
SEQUENCE FROM N.A.
RN TISSUE=Lens;
RX MEDLINE=971112991; PubMed=8943244;
RA Andley U.P., Mathur S., Griest T.A., Petrash J.M.;
RA "Cloning, expression, and chaperone-like activity of human alphaA-
RT crystallin.;"
RL J. Biol. Chem. 271:31973-31980(1996).
[3]
SEQUENCE FROM N.A.
RN MEDLINE=20289799; PubMed=10830953;
RA Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,
RA Park H.-S., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Soeda E.,
RA Ohki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K., Polley A.,
RA Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,
RA Reichwald K., Rump A., Schillhabel M., Schudy A., Zimmermann W.,
RA Rosenthal A., Kudoh J., Shibuya K., Kwasaki K., Asakawa S.,
RA Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E.,
RA Minoshima S., Shimizu N., Nordstiek G., Hornischer K., Brandt P.,
RA Scharfe M., Schoen O., Desario A., Reichelt J., Kauer G., Bloecker H.,
RA Ramser J., Beck A., Klages S., Hennig S., Riesselmann L., Dagand E.,
RA Wehrmeyer S., Borzys K., Gardiner K., Nizetic D., Francis F.,
RA Lehrach H., Reinhardt R., Yaspo M.-L.;
RT "The DNA sequence of human chromosome 21.;"
RL Nature 405:311-319(2000).
[4]
SEQUENCE
RX MEDLINE=76187952; PubMed=817940;
RA de Jong W.W., Terwindt E.C., Bloemendal H.;
RT "The amino acid sequence of the A chain of human alpha-crystallin.;"
RL FEBS Lett. 58:310-313(1975).
[5]
REVISITONS
RA Kramps J.A., de Jong W.W.;
RL Submitted (JUN-1977) to the PIR data bank.
[6]
SEQUENCE OF 1-104 FROM N.A.
RX MEDLINE=89143747; PubMed=2918909;
RA Jaworski C.J., Piatigorsky J.;
RT "A pseudo-exon in the functional human alpha A-crystallin gene.;"
RL Nature 337:752-754(1989).
[7]
SEQUENCE OF 1-63 AND 166-173 FROM N.A.
RX TISSUE=Spleen;
RX MEDLINE=87005033; PubMed=3758227;
RA McDevitt D.S., Hawkins J.W., Jaworski C.J., Piatigorsky J.;
RT "Isolation and partial characterization of the human alpha A-crystallin gene.;"
RL Exp. Eye Res. 43:285-291(1986).
[8]
SEQUENCE OF 13-21 AND 79-88
RX MEDLINE=97152999; PubMed=8999533;
RA Lampi K.J., Ma Z., Shih M., Shearer T.R., Smith J.B., Smith D.L.,
RA David L.L.;
RT "Sequence analysis of betaA3, betaB3, and betaA4 crystallins completes the identification of the major proteins in young human lens.;"
RL J. Biol. Chem. 272:2268-2275(1997).
[9]
STRUCTURE OF CARBOHYDRATE
RX MEDLINE=92112709; PubMed=1730617;
RA Roquemore E.P., Dell A., Morris H.R., Panico M., Reason A.J.,
RA Savoy L.A., Wistow G.J., Zigler J.S., Jr., Earles B.J., Hart G.W.;
RT "Vertebrate lens alpha-crystallins are modified by O-linked N-acetylglucosamine.;"
RL J. Biol. Chem. 267:555-563(1992).
[10]
DEAMIDATION
RX MEDLINE=98205214; PubMed=9543632;
RA Takemoto L.J.;
RT "Quantitation of asparagine-101 deamidation from alpha-A crystallin

RT during aging of the human lens.;"
RL Curr. Eye Res. 17:247-250(1998).
[11]
VARIANT ADCC CYS-116.
RX MEDLINE=98133928; PubMed=9467006;
RA Litt M., Kramer P., la Morticella D.M., Murphey W., Lovrien E.W.,
RA Weleber R.G.;
RT "Autosomal dominant congenital cataract associated with a missense mutation in the human alpha crystallin gene CRVA.;"
RL Hum. Mol. Genet. 7:471-474(1998)
CC -!- FUNCTION: MAY CONTRIBUTE TO THE TRANSPARENCY AND REFRACTIVE INDEX OF THE LENS.
CC -!- PTM: O-GLYCOSYLATED; CONTAINS N-ACETYLGLUCOSAMINE SIDE CHAINS.
CC -!- PTM: IN LENS, MOST DEAMIDATION OF ASN-101 OCCURS DURING THE FIRST 30 YEARS OF AGE, FOLLOWED BY A SMALL ADDITIONAL AMOUNT OF DEAMIDATION (APPROXIMATELY 5%) DURING THE NEXT APPROXIMATELY 38 YEARS, RESULTING IN A MAXIMUM OF APPROXIMATELY 50% DEAMIDATION DURING THE LIFETIME OF THE INDIVIDUAL.
CC -!- DISEASE: DEFECTS IN CRVA ARE THE CAUSE OF AUTOSOMAL DOMINANT CONGENITAL CATARACT (ADCC). ADCC IS A MAJOR ABNORMALITY OF THE EYE THAT FREQUENTLY CAUSES BLINDNESS IN INFANTS. AT LEAST A THIRD OF ALL CASES ARE FAMILIAL; ADCC APPEARS TO BE THE MOST COMMON FAMILIAL FORM IN THE WESTERN WORLD.
CC -!- SIMILARITY: BELONGS TO THE SMALL HEAT SHOCK PROTEIN (HSP20) FAMILY. STRONG TO ALPHA(B)-CRYSTALLIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL: U05569; AAA97523.1; -
DR EMBL: U66584; AAC50900.1; -
DR EMBL: X14789; CAA32891.1; -
DR EMBL: M35628; AAA52106.1; -
DR EMBL: M35629; AAA52105.1; -
DR EMBL: AP001748; BAA95535.1; -
DR PIR: A02891; CYHUA.
DR SWISS-2DPAGE; P02489; HUMAN.
DR MIM: 123580; -
DR InterPro: IPR001436; -
DR InterPro: IPR002068; -
DR InterPro: IPR003090; -
DR Pfam: PF00011; HSP20; 1.
DR Pfam: PF00525; Crystallin; 1.
DR PRINTS: PR00299; ACRYSTALLIN.
DR PROSITE: PS01031; HSP20; 1.
KW Eye lens protein; Acetylation; Glycoprotein; Disease mutation.
FT MOD_RES 1 1 ACETYLATION.
FT MOD_RES 101 101 DEAMIDATION (PARTIAL).
FT CARBOHYD 162 162 O-LINKED (GLCNAC) (BY SIMILARITY).
FT VARIANT 116 116 R -> C (IN ADCC).
FT /FTID=VAR_003819.
FT CONFLICT 45 45 S -> T (IN REF. 7).
FT CONFLICT 153 155 THA -> HT (IN REF. 1).
SQ SEQUENCE 173 AA; 19909 MW; 81804A8439837D50 CRC64;

Query Match 3.2%; Score 7; DB 1; Length 173;
Best Local Similarity 100.0%; Pred. No. 6.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 69 EEEKPTSA 75
| | | | |
DB 164 EEEKPTSA 170

RESULT 20
CRAA_RANCA
ID CRAA_RANCA STANDARD; PRT; 173 AA.

AC 091311;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE ALPHA CRYSTALLIN A CHAIN.
 GN CRYAA.
 OS Rana catesbeiana (Bull frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidea; Ranidae; Rana.
 OX NCBI_TaxID=8400;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95283564; PubMed=7763371;
 RA Lu S.F., Pan F.M., Chiou S.H.;
 RT "Sequence analysis of frog alpha-crystallin cDNA and its deduced
 RT primary structure: comparison of alpha A subunit chains among
 RT different vertebrate species.";
 RL Biochem. Biophys. Res. Commun. 210:974-981(1995).
 CC -!- FUNCTION: MAY CONTRIBUTE TO THE TRANSPARENCY AND REFRACTIVE INDEX
 CC OF THE LENS.
 CC -!- SIMILARITY: BELONGS TO THE SMALL HEAT SHOCK PROTEIN (HSP20)
 CC FAMILY. STRONG TO ALPHA(B)-CRYSTALLIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in. no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL; X85205; CAA59471.1; -
 DR InterPro; IPR001436; -
 DR InterPro; IPR002068; -
 DR InterPro; IPR003090; -
 DR Pfam; PF00011; HSP20; 1.
 DR Pfam; PF00525; crystallin; 1.
 DR PRINTS; PR00299; ACRYSTALLIN.
 DR PROSITE; PS01031; HSP20; 1.
 KW Eye lens protein; Acetylation.
 FT MOD_RES 1 ACETYLATION (BY SIMILARITY).
 SQ SEQUENCE 173 AA; 19899 MW; 31B9BD2770ABC290 CRC64;
 Query Match 3.2%; Score 7; DB 1; Length 173;
 Best Local Similarity 100.0%; Pred. No. 6.8;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 69 EEKPTSA 75
 Db 164 EEKPTSA 170
 RESULT 21
 ID CRAA_RHEAM STANDARD; PRT; 173 AA.
 AC P02505;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE ALPHA CRYSTALLIN A CHAIN.
 GN CRYAA.
 OS Rhea americana (Greater rhea) (Common rhea).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Palaeognathae; Rheiformes; Rheidae; Rhea.
 OX NCBI_TaxID=8797;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=85012654; PubMed=6482951;
 RA Stapel S.O., Leunissen J.A.M., Versteeg M., Wattel J.,
 RA de Jong W.W.;
 RT "Ratites as oldest offshoot of avian stem -- evidence from alpha-
 RT crystallin A sequences.";

RL Nature 311:257-259(1984).
 RN [2]
 RP STRUCTURE OF CARBOHYDRATE.
 RX MEDLINE=92112709; PubMed=1730617;
 RA Roquemore E.P., Dell A., Morris H.R., Panico M., Reason A.J.,
 RA Savoy L.A., Mistow G.J., Zigler J.S. Jr., Earles B.J., Hart G.W.;
 RT "Vertebrate lens alpha-crystallins are modified by O-linked N-
 RT acetylglucosamine.";
 RL J. Biol. Chem. 267:555-563(1992).
 CC -!- FUNCTION: MAY CONTRIBUTE TO THE TRANSPARENCY AND REFRACTIVE INDEX
 CC OF THE LENS.
 CC -!- PTM: O-GLYCOSYLATED; CONTAINS N-ACETYLGLUCOSAMINE SIDE CHAINS.
 CC -!- SIMILARITY: BELONGS TO THE SMALL HEAT SHOCK PROTEIN (HSP20)
 CC FAMILY. STRONG TO ALPHA(B)-CRYSTALLIN.
 CC PIR; A02908; CYEHA.
 DR GlycoSuiteDB; P02505; -
 DR InterPro; IPR001436; -
 DR InterPro; IPR002068; -
 DR InterPro; IPR003090; -
 DR Pfam; PF00011; HSP20; 1.
 DR Pfam; PF00525; crystallin; 1.
 DR PRINTS; PR00299; ACRYSTALLIN.
 DR PROSITE; PS01031; HSP20; 1.
 KW Eye lens protein; Acetylation; Glycoprotein.
 FT MOD_RES 1 ACETYLATION (PROBABLE).
 FT CARBOHYD 162 162 O-LINKED (GLCNAC) (BY SIMILARITY).
 SQ SEQUENCE 173 AA; 19832 MW; 33399BA6896DFC6C CRC64;
 Query Match 3.2%; Score 7; DB 1; Length 173;
 Best Local Similarity 100.0%; Pred. No. 6.8;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 69 EEKPTSA 75
 Db 164 EEKPTSA 170
 RESULT 22
 ID CRAA_TUPTU STANDARD; PRT; 173 AA.
 AC P02506;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE ALPHA CRYSTALLIN A CHAIN.
 GN CRYAA.
 OS Tupinambis teguixin (Common tegu).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Lepidosauria; Squamata; Scleroglossa; Scincomorpha; Teiioidea;
 OC Teiidae; Tupinambis.
 OX NCBI_TaxID=8532;
 RN [1]
 RP COMPOSITIONS OF PEPTIDES, AND PARTIAL SEQUENCE.
 RX MEDLINE=88216135; PubMed=3870872;
 RA de Jong W.W., Zweers A., Versteeg M., Dessauer H.C., Goodman M.;
 RT "Alpha-crystallin A sequences of Alligator mississippiensis and the
 RT lizard Tupinambis teguixin: molecular evolution and reptilian
 RT phylogeny.";
 RL Mol. Biol. Evol. 2:484-493(1985).
 CC -!- FUNCTION: MAY CONTRIBUTE TO THE TRANSPARENCY AND REFRACTIVE INDEX
 CC OF THE LENS.
 CC -!- SIMILARITY: BELONGS TO THE SMALL HEAT SHOCK PROTEIN (HSP20)
 CC FAMILY. STRONG TO ALPHA(B)-CRYSTALLIN.
 CC PIR; B25753; CYLZAA.
 DR InterPro; IPR001436; -
 DR InterPro; IPR002068; -
 DR InterPro; IPR003090; -
 DR Pfam; PF00011; HSP20; 1.
 DR Pfam; PF00525; crystallin; 1.
 DR PRINTS; PR00299; ACRYSTALLIN.
 DR PROSITE; PS01031; HSP20; 1.
 KW Eye lens protein; Acetylation; Glycoprotein.

FT MOD_RES 1 1 ACETYLATION.
 FT CARBOHYD 162 162 O-LINKED (GLCNAC) (BY SIMILARITY).
 SQ SEQUENCE 173 AA; 19898 MW; 3097BDD8261AAE3 CRC64;

Query Match 3.2%; Score 7; DB 1; Length 173;
 Best Local Similarity 100.0%; Pred. No. 6.8;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 69 EEKPTSA 75
 |||||
 DB 164 EEKPTSA 170

RESULT 23
 YL23 YEAST
 ID YL23 YEAST STANDARD; PRT; 259 AA.
 AC P53769.
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DE HYPOTHETICAL 29.7 KDA PROTEIN IN REC102-SFH1 INTERGENIC REGION.
 GN YLR323C OR L8543.3.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=S288C / AB972;
 RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z.,
 RA Favell A., Fulton L., Gattung S., Greco T., Kirsten J.,
 RA Kucaba T., Hallsworth K., Hawkins J., Hillier L., Jier M.,
 RA Johnson D., Johnston L., Langston Y., Latreille P., Le T.,
 RA Maridis E., Menezes S., Miller N., Nhan M., Pauley A., Peluso D.,
 RA Rifken L., Riles L., Taich A., Trevasakis E., Vignati D.,
 RA Wilcox L., Wohldman P., Vaudin M., Wilson R., Waterston R.;
 RL Submitted (Feb-1995) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
 CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.

CC -1- SIMILARITY: CONTAINS A CX(8)CX(5)CX(3)H-TYPE ZINC FINGER.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; U20618; AB64511.1; -
 DR SGD; S0004315; YLR323C.
 DR InterPro; IPR000571; -
 DR InterPro; IPR001841; -
 DR Pfam; PF000097; zf-C3HC4; 1.
 DR Pfam; PF00642; zf-CCCH; 1.
 KW Hypothetical protein; Zinc-finger; Metal-binding; DNA-binding;
 KW Nuclear protein.
 FT ZN_FING 144 163 CX(8)CX(5)CX(3)H-TYPE.
 FT ZN_FING 199 236 RING-TYPE.
 SQ SEQUENCE 259 AA; 29741 MW; 4371454CCB6F13D1 CRC64;

Query Match 3.2%; Score 7; DB 1; Length 259;
 Best Local Similarity 100.0%; Pred. No. 9.7;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 65 VASDEEK 71
 |||||
 DB 31 VASDEEK 37

RESULT 24.

MENA_ECOLI
 ID MENA_ECOLI STANDARD; PRT; 308 AA.
 AC P32166;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE 1,4-DIHYDROXY-2-NAPHTHOATE OCTAPRENYLTRANSFERASE (EC 2.5.1.-) (DHNA-
 DE OCTAPRENYLTRANSFERASE).
 GN MENA.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
 RC STRAIN=K12;
 RC MEDLINE=98241547; PubMed=95731170;
 RA Suvarna K., Stevenson D., Meganathan R., Hudspeth M.E.S.;
 RT "Menquinone (vitamin K2) biosynthesis: localization and
 RT characterization of the menA gene from Escherichia coli.";
 RL J. Bacteriol. 180:2782-2787(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RC MEDLINE=93347969; PubMed=8346018;
 RA Plunkett G. III, Burland V.D., Daniels D.L., Blattner F.R.;
 RT "Analysis of the Escherichia coli genome. III. DNA sequence of the
 RT region from 87.2 to 89.2 minutes";
 RL Nucleic Acids Res. 21:3391-3398(1993).
 CC -1- FUNCTION: CONVERSION OF 1,4-DIHYDROXY-2-NAPHTHOATE (DHNA) TO
 CC DIMETHYLMENQUINONE (DMK). ATTACHES OCTAPRENYLPYROPHOSPHATE, A
 CC MEMBRANE-BOUND 40-CARBON SIDE CHAIN TO DHNA. THE CONVERSION OF
 CC DHNA TO DMK PROCEEDS IN THREE STAGES: THE REMOVAL OF THE CARBOXYL
 CC GROUP OF DHNA AS CO₂, THE ATTACHMENT OF THE ISOPRENOID SIDE CHAIN,
 CC AND A QUINOL-TO-QUINONE OXIDATION, WHICH IS THOUGHT TO BE
 CC SPONTANEOUS.
 CC -1- PATHWAY: MENAQUINONE BIOSYNTHESIS.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE.
 CC -1- SIMILARITY: BELONGS TO THE MENA FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; U56082; AB01207.1; -
 DR EMBL; LI9201; AB03062.1; -
 DR EMBL; AE000467; AAC76912.1; -
 DR PIR; S40873; S40873.

DR EcoGene; EGI1880; menA.
 KW Menquinone biosynthesis; Transferase; Transmembrane; Inner membrane.
 FT TRANSMEM 21 41 POTENTIAL.
 FT TRANSMEM 43 63 POTENTIAL.
 FT TRANSMEM 98 118 POTENTIAL.
 FT TRANSMEM 124 144 POTENTIAL.
 FT TRANSMEM 149 169 POTENTIAL.
 FT TRANSMEM 177 197 POTENTIAL.
 FT TRANSMEM 238 258 POTENTIAL.
 FT TRANSMEM 287 307 POTENTIAL.
 SQ SEQUENCE 308 AA; 33594 MW; FECD9E85BAA40E2E CRC64;

Query Match 3.2%; Score 7; DB 1; Length 308;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 ICLSGLA 42
 |||||
 DB 106 ICLSGLA 112

DR MIM: 114890; -
DR InterPro: IPR003006; -
DR Pfam: PF00047; Ig: 3;
KW Immunoglobulin domain; Antigen; Signal; Glycoprotein; GPI-anchor.
FT SIGNAL 1 34
FT CHAIN 35 320
FT PROPEP 321 349
FT LIPID 320 320
FT DOMAIN 35 142
FT DOMAIN 160 222
FT DOMAIN 252 306
FT DISULFID 157 215
FT DISULFID 259 299
FT CARBOHYD 104 104
FT CARBOHYD 111 111
FT CARBOHYD 115 115
FT CARBOHYD 152 152
FT CARBOHYD 173 173
FT CARBOHYD 197 197
FT CARBOHYD 224 224
FT CARBOHYD 256 256
FT CARBOHYD 274 274
FT CARBOHYD 288 288
FT CARBOHYD 309 309
FT CONFLICT 114 114
FT CONFLICT 322 322
SQ SEQUENCE 349 AA; 38154 MW; AACF74DAIAC839D8 CRC64;

Query Match 3.2%; Score 7; DB 1; Length 349;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 ISNOQIT 175
Db 84 ISNOQIT 90
|||||
Search completed: May 23, 2001, 06:25:06
Job time: 200 sec

RESULT 25
CGM6_HUMAN STANDARD; PRT; 349 AA.
AC P31997; 016574; 060399;
DT 01-JUL-1993 (Rel. 26, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE CARCINOEMBRYONIC ANTIGEN CGM6 PRECURSOR (NONSPECIFIC CROSS-REACTING
DE ANTIGEN NCA-95) (ANTIGEN CD67) (CD66B ANTIGEN).
GN CGM6.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RX MEDLINE=9100398; PubMed=2208113;
RA Berling B., Kolbinger F., Grunert F., Thompson J.A.,
RA Brombacher F., Buchegger F., Vkiest S., Zimmermann W.,
RT "Cloning of a carcinoembryonic antigen gene family member expressed
RT in leukocytes of chronic myeloid leukemia patients and bone marrow.";
RL Cancer Res. 50:6534-6539(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=90165902; PubMed=2306228;
RA Arakawa F., Kuroki M., Misumi Y., Oikawa S., Nakazato H., Matsuoaka Y.,
RT "Characterization of a cDNA clone encoding a new species of the
RT nonspecific cross-reacting antigen (NCA), a member of the CEA gene
RT family.";
RL Biochem. Biophys. Res. Commun. 166:1063-1071(1990).
RN [3]
RP SEQUENCE FROM N.A.
RA Lamerdin J.E., McCready P.M., Skowronski E., Adamson A.W.,
RA Burkhardt-Schultz K., Gordon L., Kyle A., Ramirez M., Stilwagen S.,
RA Phan H., Velasco N., Ganes J., Danganan L., Poundstone P.,
RA Christensen M., Georgescu A., Avila J., Liu S., Attix C., Andreise T.,
RA Trankheim M., Amico-Keller G., Coefield J., Duarte S., Lucas S.,
RA Bruce R., Thomas P., Quan G., Kronmiller B., Arellano A.,
RA Montgomery M., Ow D., Nolan M., Trong S., Kobayashi A., Olsen A.O.,
RA Carrano A.V.;
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 1-21 FROM N.A.
RX MEDLINE=98102450; PubMed=9427723;
RA Eades-Perner A., Thompson J., van der Putten H., Zimmermann W.,
RT "Mice transgenic for the human CGM6 gene express its product, the
RT granulocyte marker CD66b, exclusively in granulocytes.";
RL Blood 91:663-672(1998).
CC 1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR.
CC 1- TISSUE SPECIFICITY: EXPRESSED IN LEUKOCYTES OF CHRONIC MYELOID
CC LEUKEMIA PATIENTS AND BONE MARROW.
CC 1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. CONTAINS 2
CC C2-LIKE AND ONE V-LIKE DOMAINS. BELONGS TO THE CARCINOEMBRYONIC
CC ANTIGEN SUBFAMILY.
CC 1- DATABASE: NAME=PROW; NOTE=CD guide CD66b entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd66b.htm".
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X52378; CAA36604.1; -
CC EMBL; M33326; AAA59914.1; -
CC EMBL; D90064; BAA14108.1; -
CC EMBL; AC004558; AAC13659.1; -
CC EMBL; Z95119; CAB08298.1; -
CC PIR; S13524; S13524.

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 06:24:41 ; Search time 47.04 Seconds
(without alignments)

550.658 Million cell updates/sec

Title: US-09-237-981-29

Perfect score: 221

Sequence: 1 MKHLVAALLVGLSLGVPQF.....INAWTAANDRWPWIOVTG 221

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 374700 seqs, 117207915 residues

Word size : 0

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 50 summaries

Database :

SPTREMBL_15:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_unclassified:*
- 13: sp_vertebrate:*
- 14: sp_virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	160	72.4	480	11	Q35474	Q35474 mus musculus
2	45	20.4	480	4	O43854	O43854 homo sapien
3	10	4.5	263	4	Q99740	Q99740 homo sapien
4	10	4.5	1193	13	Q90819	Q90819 gallus gall
5	10	4.5	1218	4	O15122	O15122 homo sapien
6	10	4.5	1218	4	O15816	O15816 homo sapien
7	10	4.5	1218	4	O14902	O14902 homo sapien
8	10	4.5	1218	11	Q9QXX0	Q9QXX0 mus musculus
9	10	4.5	1219	11	Q63722	Q63722 rattus norv
10	10	4.5	1227	4	P78504	P78504 homo sapien
11	9	4.1	1404	5	Q9VB65	Q9VB65 drosophila
12	9	4.1	2146	5	Q9VC97	Q9VC97 drosophila
13	8	3.6	63	2	Q9RTW3	Q9RTW3 deinococcus
14	8	3.6	247	2	Q9RZ15	Q9RZ15 deinococcus
15	8	3.6	252	4	O95378	O95378 homo sapien
16	8	3.6	330	4	Q9NT67	Q9NT67 homo sapien
17	8	3.6	481	10	Q9LYD5	Q9LYD5 arabidopsis
18	8	3.6	529	5	Q25058	Q25058 helioicidari
19	8	3.6	832	5	Q99108	Q99108 drosophila

20	8	3.6	1193	5	Q19617	Q19617 caenorhabdi
21	8	3.6	1212	13	O42347	O42347 gallus gall
22	8	3.6	1242	4	Q9NS15	Q9NS15 homo sapien
23	8	3.6	1253	11	O61810	O61810 mus musculus
24	8	3.6	1378	11	Q07314	Q07314 rattus norv
25	8	3.6	1395	11	Q07313	Q07313 rattus norv
26	8	3.6	1399	11	Q07280	Q07280 rattus norv
27	8	3.6	1438	11	Q07312	Q07312 rattus norv
28	8	3.6	1471	11	Q07311	Q07311 rattus norv
29	8	3.6	1578	11	Q07310	Q07310 rattus norv
30	8	3.6	1715	11	Q63375	Q63375 rattus norv
31	8	3.6	1728	11	Q63374	Q63374 rattus norv
32	8	3.6	1800	2	Q91948	Q91948 pseudomonas
33	8	3.6	2531	5	O16004	O16004 lytechinus
34	7	3.2	101	8	Q9ZYK5	Q9ZYK5 oreortyx pi
35	7	3.2	107	2	Q9JQ07	Q9JQ07 neisseria m
36	7	3.2	135	2	O53977	O53977 mycobacteri
37	7	3.2	154	11	Q9JIB7	Q9JIB7 mus musculus
38	7	3.2	172	2	O84854	O84854 chlamydia t
39	7	3.2	172	2	O9PL73	O9PL73 chlamydia m
40	7	3.2	202	11	P97766	P97766 mus musculus
41	7	3.2	280	1	Q9UYQ7	Q9UYQ7 pyrococcus
42	7	3.2	320	5	O62016	O62016 caenorhabdi
43	7	3.2	328	2	Q9XA64	Q9XA64 streptomyce
44	7	3.2	353	11	O35797	O35797 rattus norv
45	7	3.2	379	10	Q9XHR0	Q9XHR0 citrullus l
46	7	3.2	382	10	Q9SZH9	Q9SZH9 arabidopsis
47	7	3.2	437	10	O9ZTG2	O9ZTG2 vernicia fo
48	7	3.2	439	2	Q9RWH5	Q9RWH5 deinococcus
49	7	3.2	449	10	O22820	O22820 arabidopsis
50	7	3.2	453	10	Q9SYK9	Q9SYK9 arabidopsis

ALIGNMENTS

RESULT 1

ID	Q35474	PRELIMINARY;	PRT;	480 AA.
AC	Q35474;	Q35475;		
DT	01-JAN-1998	(TREMBLrel. 05, Created)		
DT	01-JAN-1998	(TREMBLrel. 05, Last sequence update)		
DE	01-OCT-2000	(TREMBLrel. 15, Last annotation update)		
DE	INTRIN-BINDING PROTEIN DEL1 PRECURSOR.			
GN	EDIL3 OR DEL1.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND CHARACTERIZATION.			
RC	TISSUE=EMBRYO;			
RX	MEDLINE=98083109; PubMed=9420328;			
RA	Hidai C., Zupancic T., Penta K., Mikhail A., Kawana M.,			
RA	Quettermous E.E., Aoka Y., Fukagawa M., Matsui Y., Platika D.,			
RA	Auerbach R., Hogan B.L.M., Shodgrass R., Quettermous T.;			
RT	*Cloning and characterization of developmental endothelial locus-1: an			
RT	embryonic endothelial cell protein that binds the alphavbeta3 integrin			
RT	receptor.";			
RL	Genes Dev. 12:21-33(1998).			
CC	-1- FUNCTION: PROMOTES ADHESION OF ENDOTHELIAL CELLS THROUGH			
CC	INTERACTION WITH THE ALPHA-V BETA3 INTEGRIN RECEPTOR. INHIBITS			
CC	FORMATION OF VASCULAR-LIKE STRUCTURES. MAY BE INVOLVED IN			
CC	REGULATION OF VASCULAR MORPHOGENESIS OR REMODELING IN EMBRYONIC			
CC	DEVELOPMENT.			
CC	-1- SUBCELLULAR LOCATION: EXTRACELLULAR MATRIX			
CC	-1- ALTERNATIVE PRODUCTS: TWO ISOFORMS; LONG ISOFORM (SHOWN HERE) AND			
CC	SHORT ISOFORM; ARE PRODUCED BY ALTERNATIVE SPLICING.			
CC	-1- TISSUE SPECIFICITY: EXPRESSED IN ANGIOBLASTS AND EARLY ENDOTHELIAL			
CC	CELLS. BY EMBRYONIC DAY 13.5, ALSO EXPRESSED IN A RESTRICTED GROUP			
CC	OF NON-ENDOTHELIAL CELLS INCLUDING CHONDROCYTES AND RETINAL			
CC	NEURONS.			
CC	-1- DEVELOPMENTAL STAGE: EXPRESSED IN THE EMBRYO FROM DAY 7. AFTER DAY			

15.5, EXPRESSION DECREASES AND DISAPPEARS COMPLETELY BY THE TIME OF BIRTH.

-1- SIMILARITY: CONTAINS 3 EGF-LIKE DOMAINS.

-1- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.

EMBL; AF031524; BAB86585.1; -

EMBL; AF031525; BAB86586.1; -

HSSP; P00740; IEDM.

MGI; 1329025; Edil3.

INTERPRO; IPR000152; -

INTERPRO; IPR000421; -

INTERPRO; IPR000561; -

INTERPRO; IPR000742; -

INTERPRO; IPR001881; -

PFAM; PF00008; EGF_3.

PFAM; PF00754; F5_F8_type_C; 2.

PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_1.

PROSITE; PS00022; EGF_1; UNKNOWN_3.

PROSITE; PS01186; EGF_CA; 1.

PROSITE; PS01187; EGF_CA; 1.

PROSITE; PS01285; FA58C_1; 2.

PROSITE; PS01286; FA58C_2; 2.

EGF-like domain; Alternative splicing; Signal; Developmental protein; Cell adhesion; Repeat; Vascularization.

KW EGF-like domain; Alternative splicing; Signal; Developmental protein; Cell adhesion; Repeat; Vascularization.

FT SIGNAL 1 16 POTENTIAL.

FT CHAIN 17 480 INTEGRIN-BINDING PROTEIN DELL.

FT DOMAIN 26 59 EGF-LIKE 1.

FT DOMAIN 78 116 EGF-LIKE 2.

FT DOMAIN 123 154 EGF-LIKE 3.

FT DOMAIN 161 311 F5/8 TYPE C 1.

FT DOMAIN 322 473 F5/8 TYPE C 2.

FT SITE 96 98 CELL ATTACHMENT SITE (POTENTIAL).

FT DISULFID 26 37 BY SIMILARITY.

FT DISULFID 31 48 BY SIMILARITY.

FT DISULFID 50 59 BY SIMILARITY.

FT DISULFID 78 89 BY SIMILARITY.

FT DISULFID 83 105 BY SIMILARITY.

FT DISULFID 107 116 BY SIMILARITY.

FT DISULFID 158 314 BY SIMILARITY.

FT DISULFID 301 476 BY SIMILARITY.

FT DISULFID 319 476 BY SIMILARITY.

FT VARSPLIC 218 221 INLQ -> VTVG (IN SHORT ISOFORM).

FT VARSPLIC 222 480 MISSING (IN SHORT ISOFORM).

FT SEQUENCE 480 AA; 53740 MW; 4CD91EF9261714D CRC64;

Query Match 72.4%; Score 160; DB 11; Length 480;

Best Local Similarity 100.0%; Pred. No. 3.3e-160;

Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKHLVAALLVGLSLGVQFGKGTICNPENCGICLSGLADDSFSCCEPGEAGPNC 60

Db 1 MKHLVAALLVGLSLGVQFGKGTICNPENCGICLSGLADDSFSCCEPGEAGPNC 60

QY 61 SVVEVASDEEKPTSGAGPCIPNPNCHNGGTCSEAYRGDTFTGYVCKCPRGNGIHCHQNI 120

Db 61 SVVEVASDEEKPTSGAGPCIPNPNCHNGGTCSEAYRGDTFTGYVCKCPRGNGIHCHQNI 120

QY 121 NCEAEPCNRNGICITDLVANTSCCEPGFEMGRNCOYKCSG 160

Db 121 NCEAEPCNRNGICITDLVANTSCCEPGFEMGRNCOYKCSG 160

RESULT 2

043854 PRELIMINARY; PRT: 480 AA.

AC 043854; 043855;

DT 01-JUN-1998 (TRENBLrel. 06, Created)

DT 01-JUN-1998 (TRENBLrel. 06, Last sequence update)

DE 01-OCT-2000 (TRENBLrel. 15, Last annotation update)

GN INTEGRIN-BINDING PROTEIN DELL PRECURSOR.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Euthera; Primates; Catarrhini; Homidae; Homo.

NCBI_TaxID=9606;

[[1]]

SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND CHARACTERIZATION.

RC TISSUE-EMBRYONIC LUNG;

RX MEDLINE=98083109; PubMed=9420328;

RA Hidai C., Zupancic T.J., Penta K., Mikhail A., Kawana M., Quettermous E.E., Aoka Y., Fukagawa M., Matsui Y., Platika D., Auerbach R., Hogan B.L.M., Snodgrass R., Quettermous T.;

"Cloning and characterization of developmental endothelial locus-1: an embryonic endothelial cell protein that binds the alphavbeta3 integrin receptor."

RL Genes Dev. 12:21-33(1998).

CC -1- FUNCTION: PROMOTES ADHESION OF ENDOTHELIAL CELLS THROUGH INTERACTION WITH THE ALPHA-V BETA3 INTEGRIN RECEPTOR. INHIBITS FORMATION OF VASCULAR-LIKE STRUCTURES. MAY BE INVOLVED IN REGULATION OF VASCULAR MORPHOGENESIS OF REMODELING IN EMBRYONIC DEVELOPMENT.

CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR MATRIX.

CC -1- ALTERNATIVE PRODUCTS: TWO ISOFORMS; LONG ISOFORM (SHOWN HERE) AND SHORT ISOFORM; ARE PRODUCED BY ALTERNATIVE SPLICING.

CC -1- SIMILARITY: CONTAINS 3 EGF-LIKE DOMAINS.

CC -1- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.

EMBL; U70312; AAC02648.1; -

EMBL; U70313; AAC02649.1; -

HSSP; P00740; IIXA.

INTERPRO; IPR000152; -

INTERPRO; IPR000421; -

INTERPRO; IPR000561; -

INTERPRO; IPR000742; -

INTERPRO; IPR001881; -

PFAM; PF00008; EGF_3.

PFAM; PF00754; F5_F8_type_C; 2.

PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_1.

PROSITE; PS00022; EGF_1; UNKNOWN_2.

PROSITE; PS01186; EGF_CA; 1.

PROSITE; PS01187; EGF_CA; 1.

PROSITE; PS01285; FA58C_1; 2.

PROSITE; PS01286; FA58C_2; 2.

EGF-like domain; Alternative splicing; Signal; Developmental protein; Cell adhesion; Repeat; Vascularization.

KW EGF-like domain; Alternative splicing; Signal; Developmental protein; Cell adhesion; Repeat; Vascularization.

FT SIGNAL 1 16 POTENTIAL.

FT CHAIN 17 480 INTEGRIN-BINDING PROTEIN DELL.

FT DOMAIN 26 59 EGF-LIKE 1.

FT DOMAIN 78 116 EGF-LIKE 2.

FT DOMAIN 123 154 EGF-LIKE 3.

FT DOMAIN 161 311 F5/8 TYPE C 1.

FT DOMAIN 322 473 F5/8 TYPE C 2.

FT SITE 96 98 CELL ATTACHMENT SITE (POTENTIAL).

FT DISULFID 26 37 BY SIMILARITY.

FT DISULFID 31 48 BY SIMILARITY.

FT DISULFID 50 59 BY SIMILARITY.

FT DISULFID 78 89 BY SIMILARITY.

FT DISULFID 83 105 BY SIMILARITY.

FT DISULFID 107 116 BY SIMILARITY.

FT DISULFID 158 314 BY SIMILARITY.

FT DISULFID 301 305 BY SIMILARITY.

FT DISULFID 319 476 BY SIMILARITY.

FT VARSPLIC 66 66 A -> G (IN SHORT ISOFORM).

FT VARSPLIC 67 76 MISSING (IN SHORT ISOFORM).

FT SEQUENCE 480 AA; 53765 MW; F7171E23A309FD48 CRC64;

Query Match 20.4%; Score 45; DB 4; Length 480;

Best Local Similarity 100.0%; Pred. No. 5.9e-39;

Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 PNCHNGTCEISEAYRGDTFTGYVCKCPRGNGIHCHQNI 124

Db 80 PNCHNGTCEISEAYRGDTFTGYVCKCPRGNGIHCHQNI 124

RESULT 3

Q99740 PRELIMINARY; PRT; 263 AA.
 AC Q99740;
 DT 01-MAY-1997 (TrEMBLrel. 03, Created)
 DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
 DE SOLUBLE PROTEIN JAGGED (FRAGMENT).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-UMBILICAL VEIN;
 RX MEDLINE=95211842; PubMed=7697721;
 RA Lindsell C.E., Shawber C.J., Boulter J., Weinmaster G.;
 RA "Jagged: a mammalian ligand that activates Notch1.";
 RL Cell 80:909-917(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE-UMBILICAL VEIN;
 RX MEDLINE=97115768; PubMed=8955070;
 RA Zimin A.B., Pepper M.S., McMahon G., Nguyen F., Montesano R.,
 RA MacIag T.;
 RA "An antisense oligonucleotide to the notch ligand jagged enhances
 RT fibroblast growth factor-induced angiogenesis in vitro.";
 RL J. Biol. Chem. 271:32499-32502(1996).
 DR EMBL; U77914; AAC50909.1; -;
 DR HSSP; P00743; ICCP;
 DR INTERPRO; IPR000083; -;
 DR INTERPRO; IPR000152; -;
 DR INTERPRO; IPR000561; -;
 DR INTERPRO; IPR000742; -;
 DR INTERPRO; IPR001881; -;
 DR INTERPRO; IPR002049; -;
 DR PFAM; PF00008; EGF_3;
 DR PRINTS; PR00011; EGFLAMININ.
 DR PROSITE; PS00012; FNTYPEI.
 DR PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_2.
 DR PROSITE; PS00022; EGF_1; UNKNOWN_3.
 DR PROSITE; PS01186; EGF_2; 2.
 DR PROSITE; PS01187; EGF_CA; 2.
 DR Glycoprotein; EGF-like domain.
 FT NON_TER 1
 SQ SEQUENCE 263 AA; 28405 MW; E77235CBB360B620 CRC64;
 Query Match 4.5%; Score 10; DB 4; Length 263;
 Best Local Similarity 100.0%; Pred. No. 0.027;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 80 PNPCHNGGTC 89
 Db 8 PNPCHNGGTC 17
 RESULT 4
 ID Q90819 PRELIMINARY; PRT; 1193 AA.
 AC Q90819;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
 DE C-SERATE-1 PROTEIN (FRAGMENT).
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-SPINAL CORD;
 RX MEDLINE=96175595; PubMed=8631496;

Myat A., Henrique D., Ish-Horowicz D., Lewis J.;
 RT "A chick homologue of Serrate and its relationship with Notch and
 Dev. Biol. 174:233-247(1996).
 DR EMBL; X95283; CAA64604.1; -;
 DR HSSP; P00740; IEDM.
 DR INTERPRO; IPR000083; -;
 DR INTERPRO; IPR000152; -;
 DR INTERPRO; IPR000561; -;
 DR INTERPRO; IPR000742; -;
 DR INTERPRO; IPR001010; -;
 DR INTERPRO; IPR001438; -;
 DR INTERPRO; IPR001774; -;
 DR INTERPRO; IPR001881; -;
 DR INTERPRO; IPR001947; -;
 DR INTERPRO; IPR002049; -;
 DR PFAM; PF00008; EGF; 14.
 DR PFAM; PF01414; DSL; 1.
 DR PRINTS; PR00010; EGF_BLOOD.
 DR PRINTS; PR00011; EGFLAMININ.
 DR PRINTS; PR00012; FNTYPEI.
 DR PRINTS; PR00286; CHARYBDOTOXIN.
 DR PRINTS; PR00287; THIONIN.
 DR PROSITE; PS00010; ASX_HYDROXYL; 10.
 DR PROSITE; PS00022; EGF_1; UNKNOWN_16.
 DR PROSITE; PS01186; EGF_2; 12.
 DR PROSITE; PS01187; EGF_CA; 8.
 KW Glycoprotein; EGF-like domain.
 FT NON_TER 1
 SQ SEQUENCE 1193 AA; 131039 MW; 24E94D09EB987935 CRC64;
 Query Match 4.5%; Score 10; DB 13; Length 1193;
 Best Local Similarity 100.0%; Pred. No. 0.11;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 80 PNPCHNGGTC 89
 Db 724 PNPCHNGGTC 733
 RESULT 5
 ID 015122 PRELIMINARY; PRT; 1218 AA.
 AC 015122;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
 DE JAGGED1.
 GN JAG1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Oda T., Elkhouloun A.G., Pike B.L., Okajima K., Krantz I.D., Genin A.,
 RA Piccoli D.A., Meltzer P.S., Spinner N.B., Collins F.S.,
 RA Chandrasekharappa S.C.;
 RL Nat. Genet. 0:0-0(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97422615; PubMed=9268641;
 RA Oda T., Elkhouloun A.G., Meltzer P.S., Chandrasekharappa S.C.;
 RT "Identification and cloning of the human homolog (JAG1) of the rat
 RL Jagged1 gene from the Alagille syndrome critical region at 20p12.";
 DR EMBL; AF003837; AAC51731.1; -;
 DR HSSP; P00740; IIXA.
 DR INTERPRO; IPR000152; -;
 DR INTERPRO; IPR000561; -;
 DR INTERPRO; IPR000742; -;
 DR INTERPRO; IPR001438; -;

DR INTERPRO: IPR001774; -
 DR INTERPRO: IPR001881; -
 DR PFAM: PF00008; EGF_14;
 DR PFAM: PF01414; DSL; 1.
 DR PRINTS: PR00010; EGFLOOD.
 DR PROSITE: PS00010; ASX_HYDROXYL; UNKNOWN_10.
 DR PROSITE: PS00022; EGF_1; UNKNOWN_16.
 DR PROSITE: PS01186; EGF_2; 12.
 DR PROSITE: PS01187; EGF_CA; 8.
 KW Glycoprotein; EGF-like domain.
 SQ SEQUENCE 1218 AA; 133858 MW; F34EE15FE265377C CRC64;

Query Match 4.5%; Score 10; DB 4; Length 1218;
 Best Local Similarity 100.0%; Pred. No. 0.11;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 PNPCHNGGTC 89
 |||||
 DB 750 PNPCHNGGTC 759

RESULT 6
 Q15816 PRELIMINARY; PRT; 1218 AA.
 AC Q15816;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
 DE TRANSMEMBRANE PROTEIN JAGGED 1.
 GN HJ1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95211842; PubMed=7697721;
 RA Lindsell C.E., Shawber C.J., Boulter J., Weinmaster G.;
 RT "Jagged: a mammalian ligand that activates Notch1.";
 RL Cell 80:909-917(1995).
 RN [2]
 RP SEQUENCE FROM N.A.

RA Gray G.E., Mann R.S., Mitsiadis E., Henrique D., Caracangiu M.,
 RA Ish-Horowicz D., Artavanis-Tsakonas S.;
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U61276; AAB39007.1; -
 DR HSSP; P00740; 11XA.
 DR INTERPRO: IPR000152; -
 DR INTERPRO: IPR000561; -
 DR INTERPRO: IPR000742; -
 DR INTERPRO: IPR001438; -
 DR INTERPRO: IPR001774; -
 DR INTERPRO: IPR001881; -
 DR PFAM: PF00008; EGF_14;
 DR PFAM: PF01414; DSL; 1.
 DR PRINTS: PR00010; EGFLOOD.
 DR PROSITE: PS00010; ASX_HYDROXYL; UNKNOWN_10.
 DR PROSITE: PS00022; EGF_1; UNKNOWN_16.
 DR PROSITE: PS01186; EGF_2; 12.
 DR PROSITE: PS01187; EGF_CA; 8.
 KW Transmembrane; Glycoprotein; EGF-like domain.
 SQ SEQUENCE 1218 AA; 133738 MW; 100EDC4E40AD070A CRC64;

Query Match 4.5%; Score 10; DB 4; Length 1218;
 Best Local Similarity 100.0%; Pred. No. 0.11;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 PNPCHNGGTC 89
 |||||
 DB 750 PNPCHNGGTC 759

RESULT 7
 Q14902 PRELIMINARY; PRT; 1218 AA.
 ID Q14902;
 AC Q14902;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
 DE TRANSMEMBRANE PROTEIN JAGGED 1.
 GN HJ1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95211842; PubMed=7697721;
 RA Lindsell C.E., Shawber C.J., Boulter J., Weinmaster G.;
 RT "Jagged: a mammalian ligand that activates Notch1.";
 RL Cell 80:909-917(1995).
 RN [2]
 RP SEQUENCE FROM N.A.

RA Bash J., Zong W.-X., Gelinas C.;
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF028593; AAB84053.1; -
 DR HSSP; P00740; 11XA.
 DR INTERPRO: IPR000152; -
 DR INTERPRO: IPR000561; -
 DR INTERPRO: IPR000742; -
 DR INTERPRO: IPR001438; -
 DR INTERPRO: IPR001774; -
 DR INTERPRO: IPR001881; -
 DR PFAM: PF00008; EGF_14;
 DR PFAM: PF01414; DSL; 1.
 DR PRINTS: PR00010; EGFLOOD.
 DR PROSITE: PS00010; ASX_HYDROXYL; UNKNOWN_10.
 DR PROSITE: PS00022; EGF_1; UNKNOWN_16.
 DR PROSITE: PS01186; EGF_2; 12.
 DR PROSITE: PS01187; EGF_CA; 8.
 KW Transmembrane; Glycoprotein; EGF-like domain.
 SQ SEQUENCE 1218 AA; 133797 MW; F36E9F8F64DF162 CRC64;

Query Match 4.5%; Score 10; DB 4; Length 1218;
 Best Local Similarity 100.0%; Pred. No. 0.11;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 PNPCHNGGTC 89
 |||||
 DB 750 PNPCHNGGTC 759

RESULT 8
 Q9QXX0 PRELIMINARY; PRT; 1218 AA.
 ID Q9QXX0;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE JAGGED1.
 GN JAG1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SWISS WEBSTER/NIH;
 RA Shimizu K., Chiba S., Kumano K., Hosoya N., Takahashi T., Kanda Y.,
 RA Hamada Y., Yazaki Y., Hirai H.;
 RT "Mouse Jagged1 physically interacts with Notch2 and other Notch
 RT receptors: assessment by quantitative methods.";
 RL J. Biol. Chem. 274:32961-32969(1999).

```
DR EMBL; AF171092; AAF1505.1; -.
DR HSP; P00740; 1IXA.
DR MGD; MGI:1095416; Jag1.
DR INTERPRO; IPR000152; -.
DR INTERPRO; IPR000561; -.
DR INTERPRO; IPR000742; -.
DR INTERPRO; IPR001438; -.
DR INTERPRO; IPR001774; -.
DR INTERPRO; IPR001881; -.
DR INTERPRO; IPR002049; -.
DR PFAM; PF00008; EGF; 14.
DR PFAM; PF01414; DSL; 1.
DR PRINTS; PR00010; EGFBL00D.
DR PRINTS; PR00011; EGFAMININ.
DR PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_10.
DR PROSITE; PS00022; EGF_1; UNKNOWN_16.
DR PROSITE; PS01186; EGF_2; 12.
DR PROSITE; PS01187; EGF_CA; 8.
DR PROSITE; PS01187; EGF_CA; 8.
SQ SEQUENCE 1218 AA; 134163 MW; 77739F8928BB793C CRC64;

Query Match 4.5%; Score 10; DB 11; Length 1218;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 PNPCHNGGTC 89
Db 750 PNPCHNGGTC 759

RESULT 9
O63722 PRELIMINARY; PRT; 1219 AA.
AC O63722; P70640; -.
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE JAGGED PROTEIN.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SCIATIC NERVE;
RX MEDLINE=95211842; PubMed=7697721;
RA Lindsell C.E., Shawber C.J., Boulter J., Weinmaster G.;
RT "Jagged: a mammalian ligand that activates Notch1.";
RL Cell 80:909-917(1995).
DR EMBL; L38483; AAB06509.1; -.
DR HSP; P00740; 1IXA.
DR INTERPRO; IPR000152; -.
DR INTERPRO; IPR000561; -.
DR INTERPRO; IPR000742; -.
DR INTERPRO; IPR001438; -.
DR INTERPRO; IPR001774; -.
DR INTERPRO; IPR001881; -.
DR PFAM; PF00008; EGF; 14.
DR PFAM; PF01414; DSL; 1.
DR PRINTS; PR00010; EGFBL00D.
DR PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_10.
DR PROSITE; PS00022; EGF_1; UNKNOWN_16.
DR PROSITE; PS01186; EGF_2; 12.
DR PROSITE; PS01187; EGF_CA; 8.
KW Glycoprotein; EGF-like domain.
SQ SEQUENCE 1219 AA; 134325 MW; 65D4CFC238A0E204 CRC64;

Query Match 4.5%; Score 10; DB 11; Length 1219;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 PNPCHNGGTC 89
```

```
Db 750 PNPCHNGGTC 759
|||||
RESULT 10
P78504 PRELIMINARY; PRT; 1227 AA.
ID P78504;
AC P78504;
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE JAGGED 1 (TRANSMEMBRANE PROTEIN JAGGED).
GN JG1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95211842; PubMed=7697721;
RA Lindsell C.E., Shawber C.J., Boulter J., Weinmaster G.;
RT "Jagged: a mammalian ligand that activates Notch1.";
RL Cell 80:909-917(1995).
RN [2]
RP SEQUENCE FROM N.A.
RA Li L., Deng Y., Banta A.B., Hood L.;
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 14-1227 FROM N.A.
RX MEDLINE=97115768; PubMed=8955070;
RA Zimin A.B., Pepper M.S., McMahon G., Nguyen F., Montesano R.,
RA Maciag T.;
RT "An antisense oligonucleotide to the notch ligand jagged enhances
fibroblast growth factor-induced angiogenesis in vitro.";
RL J. Biol. Chem. 271:32499-32502(1996).
RN [4]
RP REVISIONS TO 14-1227.
RA Zimin A.B., Nguyen F., Maciag T.;
RA Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U73936; AAC52020.1; -.
DR EMBL; U77720; AAC51323.1; -.
DR HSP; P00740; 1IXA.
DR INTERPRO; IPR000152; -.
DR INTERPRO; IPR000561; -.
DR INTERPRO; IPR000742; -.
DR INTERPRO; IPR001438; -.
DR INTERPRO; IPR001774; -.
DR INTERPRO; IPR001881; -.
DR PFAM; PF00008; EGF; 14.
DR PFAM; PF01414; DSL; 1.
DR PRINTS; PR00010; EGFBL00D.
DR PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_10.
DR PROSITE; PS00022; EGF_1; UNKNOWN_16.
DR PROSITE; PS01186; EGF_2; 12.
DR PROSITE; PS01187; EGF_CA; 8.
KW Transmembrane.
FT CONFLICT 1187 1227
FT FT
FT REF. 1 AND 2.
SQ SEQUENCE 1227 AA; 134770 MW; 98EB7D21CB56C15 CRC64;

Query Match 4.5%; Score 10; DB 4; Length 1227;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 PNPCHNGGTC 89
|||||
Db 750 PNPCHNGGTC 759

RESULT 11
Q9VB65
```

ID Q9VB65 PRELIMINARY; PRT; 1404 AA.
 AC Q9VB65;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE SER PROTEIN.
 GN SER.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbán K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Glöckle A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Liao X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Wu D., Yang S., Yao Q.A.,
 RA Williams S.M., Woodage T., Worley K.C., Zhang G., Zhao Q., Zheng L.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 DR EMBL; AE003759; AAC56678.1; -;
 DR HSP; P00743; ICCF.
 DR FLYBASE; FBgn0004197; Ser.
 DR INTERPRO; IPR000152; -;
 DR INTERPRO; IPR000255; -;
 DR INTERPRO; IPR000515; -;
 DR INTERPRO; IPR000561; -;
 DR INTERPRO; IPR000742; -;
 DR INTERPRO; IPR001438; -;
 DR INTERPRO; IPR001774; -;
 DR INTERPRO; IPR001881; -;
 DR INTERPRO; IPR002049; -;
 DR PRAM; PF00008; EGF; 11.
 DR PRAM; PF01414; DSL; 1.
 DR PRINTS; PR00010; EGFBL00.
 DR PRINTS; PR00011; EGFBLMIN.
 DR PROSITE; PS00010; ASX_HYDROXYL; 7.
 DR PROSITE; PS00012; PHOSPHOPANTETHEINE; UNKNOWN_1.
 DR PROSITE; PS00022; EGF_1; 14.

DR PROSITE; PS00402; BPD_TRANS_INN_MEMBER; UNKNOWN_1.
 DR PROSITE; PS01186; EGF_2; 8.
 DR PROSITE; PS01187; EGF_CA; 5.
 SQ SEQUENCE 1404 AA; 150342 MW; E988604001DAAC84 CRC64;

 Query Match 4.1%; Score 9; DB 5; Length 1404;
 Best Local Similarity 100.0%; Pred. No. 1.4;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 82 PCHNGTCE 90
 DB 883 PCHNGTCE 891
 |||||
 |||||

 RESULT 12
 Q9VC97 PRELIMINARY; PRT; 2146 AA.
 AC Q9VC97;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE CRB PROTEIN.
 GN CRB.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbán K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Glöckle A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Liao X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Wu D., Yang S., Yao Q.A.,
 RA Williams S.M., Woodage T., Worley K.C., Zhang G., Zhao Q., Zheng L.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 DR EMBL; AE003747; AAF56276.1; -;
 DR HSP; P00740; IIXA.

DR FLYBASE; FBgn0000368; crb.
DR INTERPRO; IPR000152; -
DR INTERPRO; IPR000561; -
DR INTERPRO; IPR000742; -
DR INTERPRO; IPR001438; -
DR INTERPRO; IPR001791; -
DR INTERPRO; IPR001861; -
DR INTERPRO; IPR002049; -
DR PFAM; PF00008; EGF; 27.
DR PFAM; PF00054; laminin_G; 3.
DR PRINTS; PR00010; EGFBLD.
DR PRINTS; PR00011; EGFBLMIN.
DR PROSITE; PS00010; ASX_HYDROXYL; 16.
DR PROSITE; PS00022; EGF_1; 25.
DR PROSITE; PS01186; EGF_2; 17.
DR PROSITE; PS01187; EGF_CA; 13.
SQ SEQUENCE 2146 AA; 233570 MW; 8E23B9E32B761115 CRC64;

Query Match 4.1%; Score 9; DB 5; Length 2146;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 48 CECPEGFAG 56
| | | | | | | | | |
DB 2056 CECPEGFAG 2064

RESULT 13
Q9RTW3 PRELIMINARY; PRT; 63 AA.
AC Q9RTW3;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2000 (Tremblrel. 14, Last annotation update)
DE HYPOTHETICAL 7.0 KDA PROTEIN.
GN DR1640.
OS Deinococcus radiodurans.
OC Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zalewski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome Sequence of the Radioresistant Bacterium Deinococcus
RT radiodurans R1.";
RL Science 286:1571-1577(1999).
DR EMBL; AE001862; AAF1201.1; -
DR TIGR; DR1640; -
DR INTERPRO; IPR000515; -
DR PFAM; PF00528; BPD_transp; 1.
SQ SEQUENCE 247 AA; 26293 MW; 80E694BE82938282 CRC64;

Query Match 3.6%; Score 8; DB 2; Length 247;
Best Local Similarity 100.0%; Pred. No. 3.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 LLVGLSLG 16
| | | | | | | | | |
DB 67 LLVGLSLG 74

RESULT 15
Q95378 PRELIMINARY; PRT; 252 AA.
AC Q95378;
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DT 01-JUN-2000 (Tremblrel. 14, Last annotation update)
DE NEUREXIN III-ALPHA (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Rowen L., Madan A., Qin S., Ratcliffe A., Shaffer T., Abbasi N.,
RA Dickhoff R., James R., Loretz C., Madan A., Dors M., Hood L.;
RL "Sequence of chromosome 14 from 14q24.3-14q32.";
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF099810; AAC68909.1; -
DR INTERPRO; IPR000991; -
DR INTERPRO; IPR001791; -
DR PFAM; PF00054; laminin_G; 1.
DR PROSITE; PS00442; GATASE_TYPE_I; UNKNOWN_1.
FT NON_TER 252 252
SQ SEQUENCE 252 AA; 27508 MW; F805A8A013C836D6 CRC64;

Query Match 3.6%; Score 8; DB 4; Length 252;
Best Local Similarity 100.0%; Pred. No. 3.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 PCENGIC 37
| | | | | | | | | |
DB 206 PCENGIC 213

```

RESULT 16
Q9NT67 ID Q9NT67 PRELIMINARY; PRT; 330 AA.
AC Q9NT67;
DT 01-OCT-2000 (TREMELrel. 15, Created)
DT 01-OCT-2000 (TREMELrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMELrel. 15, Last annotation update)
DE HYPOTHETICAL 36.5 KDA PROTEIN (FRAGMENT).
GN DKF2P761L191.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE=AMYGDALA;
RC Ottenwaelder B., Obermaier B., Mewes H.W., Gassenhuber J., Wiemann S.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL137504; CAB70776.1; -.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 330 AA; 36465 MW; CAA0BF2DB9C7552D CRC64;

Query Match 3.6%; Score 8; DB 4; Length 330;
Best Local Similarity 100.0%; Pred. No. 4.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 82 PCHNGGTC 89
Db 73 PCHNGGTC 80

RESULT 17
Q9LYD5 ID Q9LYD5 PRELIMINARY; PRT; 481 AA.
AC Q9LYD5;
DT 01-OCT-2000 (TREMELrel. 15, Created)
DT 01-OCT-2000 (TREMELrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMELrel. 15, Last annotation update)
DE PUTATIVE OLIGOPEPTIDE TRANSPORTER PROTEIN.
GN F15N18.160.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Hilbert H., Braun M., Holzer E., Brandt A., Duesterhoeft A.,
RA Bancroft I., Mewes H.W., Rudd S., Lemcke K., Mayer K.F.X.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL163815; CAB87717.1; -.
SQ SEQUENCE 481 AA; 53365 MW; A6D3F18A110BB5F5 CRC64;

Query Match 3.6%; Score 8; DB 10; Length 481;
Best Local Similarity 100.0%; Pred. No. 6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 WLLVGLSL 15
Db 456 WLLVGLSL 463

RESULT 18
Q25058 ID Q25058 PRELIMINARY; PRT; 529 AA.

```

```

Q25058;
AC Q25058;
DT 01-NOV-1996 (TREMELrel. 01, Created)
DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)
DT 01-OCT-2000 (TREMELrel. 15, Last annotation update)
DE FIBROPELIN IA (FRAGMENT).
OS Helicoidaris erythrogramma (Sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidea; Euechinozoa; Echinoida; Echinacea; Echinommatidae;
OC Helicoidaris.
OX NCBI_TaxID=7634;
RN [1]
RP SEQUENCE FROM N.A.
RA Bisgrove B.W.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; L33861; AAA29995.1; -.
DR HSSP; P08709; 1BF9.
DR INTERPRO; IPR000083; -.
DR INTERPRO; IPR000088; -.
DR INTERPRO; IPR000152; -.
DR INTERPRO; IPR000561; -.
DR INTERPRO; IPR000742; -.
DR INTERPRO; IPR001010; -.
DR INTERPRO; IPR001438; -.
DR INTERPRO; IPR001881; -.
DR INTERPRO; IPR001947; -.
DR INTERPRO; IPR002049; -.
DR PFAM; PF00008; EGF_10;
DR PFAM; PF01382; Avidin; 1.
DR PRINTS; PR00010; EGFELOOD.
DR PRINTS; PR00011; EGFELAMININ.
DR PRINTS; PR00012; FNTYPEI.
DR PRINTS; PR00286; CHARYBOTOXIN.
DR PRINTS; PR00287; THIONIN.
DR PRINTS; PR00709; AVIDIN.
DR PROSITE; PS00010; ASX HYDROXYL; UNKNOWN_8.
DR PROSITE; PS00022; EGF_1; UNKNOWN_11.
DR PROSITE; PS00577; AVIDIN; 1.
DR PROSITE; PS01186; EGF_2; 10.
DR PROSITE; PS01187; EGF_CA; 7.
KW Glycoprotein; EGF-like domain.
FT NON_TER 1
SQ SEQUENCE 529 AA; 55543 MW; D4AE958FCF9ACB5A CRC64;

Query Match 3.6%; Score 8; DB 5; Length 529;
Best Local Similarity 100.0%; Pred. No. 6.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 PCENGGIC 37
Db 181 PCENGGIC 188

RESULT 19
Q99108 ID Q99108 PRELIMINARY; PRT; 832 AA.
AC Q99108;
DT 01-NOV-1996 (TREMELrel. 01, Created)
DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)
DT 01-JUN-2000 (TREMELrel. 14, Last annotation update)
DE NEUROGENIC LOCUS DELTA PROTEIN PRECURSOR (VERSION 2).
GN DL.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OREGON-R; TISSUE=EMBRYO;
RX MEDLINE=89196890; PubMed=3149249;
RA Kopcynski C.C., Alton A.K., Fechtel K., Koeh P.J., Muskavitch M.A.T.;
RT "Delta, a Drosophila neurogenic gene, is transcriptionally complex and

```

RT encodes a protein related to blood coagulation factors and epidermal
 RL growth factor of vertebrates.";
 RL Genes Dev. 2:1723-1735(1988).
 [2]
 RN SEQUENCE OF 422-621 FROM N.A.
 RP MEDLINE=87218537; PubMed=3107986;
 RX Knust E., Dietrich U., Tepass U., Bremer K.A., Weigel D., Vaessin H.,
 RA Campos-Ortega J.A.;
 RA "EGF homologous sequences encoded in the genome of Drosophila
 RT melanogaster, and their relation to neurogenic genes.";
 RL EMBO J. 6:761-766(1987).
 [3]
 RN PATTERN OF TRANSCRIPTION.
 RP MEDLINE=91209246; PubMed=2128477;
 RX Haenlin M., Kramatschek B., Campos-Ortega J.A.;
 RA "The pattern of transcription of the neurogenic gene Delta of
 RT Drosophila melanogaster.";
 RL Development 110:905-914(1990).
 CC -!- FUNCTION: ESSENTIAL FOR PROPER DIFFERENTIATION OF ECTODERM. DL IS
 CC REQUIRED FOR THE CORRECT SEPARATION OF NEURAL AND EPIDERMAL CELL
 CC LINESAGES.
 CC -!- MISCELLANEOUS: SEPARATION OF NEUROBLASTS FROM THE ECTODERM INTO
 CC THE INNER PART OF EMBRYO IS ONE OF THE FIRST STEPS OF CNS
 CC DEVELOPMENT IN INSECTS, THIS PROCESS IS UNDER CONTROL OF THE
 CC NEUROGENIC GENES.
 CC -!- MISCELLANEOUS: NOTCH AND SERRATE MAY INTERACT AT THE PROTEIN MAY
 CC LEVEL, IT IS CONCEIVABLE THAT THE SERRATE AND DELTA PROTEINS MAY
 CC COMPETE FOR BINDING WITH THE NOTCH PROTEIN.
 CC -!- SIMILARITY: THE PROTEIN INCLUDES 9 EGF-LIKE REPEATS.
 CC -!- SIMILARITY: TO DROSOPHILA SERRATE PROTEIN (AC P18168), AND
 CC VERTEBRATE BLOOD COAGULATION FACTOR IX.
 DR EMBL; Y00222; CAA68369.1; -;
 DR HSP; P00740; 11XA.
 DR FLYBASE; FBgn0000463; DL.
 DR INTERPRO; IPR000152; -;
 DR INTERPRO; IPR000561; -;
 DR INTERPRO; IPR000742; -;
 DR INTERPRO; IPR000847; -;
 DR INTERPRO; IPR001774; -;
 DR INTERPRO; IPR001881; -;
 DR PFAM; PF00008; EGF; 9.
 DR PFAM; PF01414; DSL; 1.
 DR PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_3.
 DR PROSITE; PS00022; EGF_1; UNKNOWN_9.
 DR PROSITE; PS00044; HTH_LYSR_FAMILY; UNKNOWN_1.
 DR PROSITE; PS01186; EGF 2; 9.
 DR PROSITE; PS01187; EGF CA; 2.
 KW Differentiation; Neurogenesis; Repeat; Transmembrane; EGF-like domain;
 KW Glycoprotein; Signal.
 FT SIGNAL 1 25 POTENTIAL.
 FT CHAIN 26 832 DELTA PROTEIN.
 FT DOMAIN 26 595 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 217 566 9 EGF-TYPE REPEATS.
 FT TRANSMEM 569 617 POTENTIAL.
 FT DOMAIN 618 832 INTRACELLULAR (POTENTIAL).
 FT REPEAT 217 256 EGF-LIKE 1.
 FT REPEAT 257 290 EGF-LIKE 2.
 FT REPEAT 291 330 EGF-LIKE 3.
 FT REPEAT 331 373 EGF-LIKE 4.
 FT REPEAT 374 417 EGF-LIKE 5.
 FT REPEAT 418 452 EGF-LIKE 6.
 FT REPEAT 453 490 EGF-LIKE 7.
 FT REPEAT 491 528 EGF-LIKE 8.
 FT REPEAT 529 566 EGF-LIKE 9.
 FT CARBOHYD 98 98 POTENTIAL.
 FT CARBOHYD 137 137 POTENTIAL.
 FT CARBOHYD 167 167 POTENTIAL.
 FT CARBOHYD 421 421 POTENTIAL.
 FT CARBOHYD 649 649 POTENTIAL.
 FT CONFLICT 437 438 GK -> ET (IN REF. 2).
 FT CONFLICT 443 443 A -> S (IN REF. 2).
 FT CONFLICT 459 459 G -> A (IN REF. 2).
 FT CONFLICT 490 490 S -> T (IN REF. 2).

FT CONFLICT 591 591 T -> A (IN REF. 2).
 SQ SEQUENCE 832 AA; 88813 MW; 5478A389D9F29770 CRC64;
 Query Match 3.6%; Score 8; DB 5; Length 832;
 Best Local Similarity 100.0%; Pred. No. 9.8;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 82 PCHNGGTC 89
 Db 537 PCHNGGTC 544
 RESULT 20
 Q19617
 ID Q19617 PRELIMINARY; PRT; 1193 AA.
 AC Q19617;
 DT 01-NOV-1996 (T-EMBLrel. 01, Created)
 DT 01-AUG-1998 (T-EMBLrel. 07, Last sequence update)
 DT 01-OCT-2000 (T-EMBLrel. 15, Last annotation update)
 DE F20B10.1 PROTEIN.
 DE F20B10.1.
 GN F20B10.1.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 OX [1]
 RN SEQUENCE FROM N.A.
 RA Percy C.;
 RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94150718; PubMed=7906398;
 RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
 RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkneen R.,
 RA Smalton N., Smith A., Sonhammer E., Staden R., Sulston J.,
 RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
 RA Watson A., Weinstock L., Wilkison-Sproat J., Wohlman P.;
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RL elegans.";
 RL Nature 368:32-38(1994).
 DR EMBL; Z69636; CAA93465.1; -;
 DR HSP; P01132; 1EGF.
 DR INTERPRO; IPR000152; -;
 DR INTERPRO; IPR000561; -;
 DR INTERPRO; IPR000742; -;
 DR INTERPRO; IPR001791; -;
 DR PFAM; PF00008; EGF; 2.
 DR PFAM; PF00054; laminin.G; 1.
 DR PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_1.
 DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
 DR PROSITE; PS01186; EGF_2; 1.
 SQ SEQUENCE 1193 AA; 132494 MW; FEB4D20DEC53E22A CRC64;
 Query Match 3.6%; Score 8; DB 5; Length 1193;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 28 PNPCEGG 35
 Db 463 PNPCEGG 470
 RESULT 21
 O42347
 ID O42347 PRELIMINARY; PRT; 1212 AA.
 AC O42347;

DT 01-JAN-1998 (TReMBLrel. 05, Created)
 DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
 DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
 DE C-SERRATE-2 (FRAGMENT)
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97184054; PubMed=9032014;
 RA Hayashi H., Mochii M., Kodama R., Hamada Y., Mizuno N., Eguchi G.,
 RA Tachi C.;
 RT "Isolation of a novel chick homolog of Serrate and its coexpression
 with C-Notch-1 in chick development.";
 RL Int. J. Dev. Biol. 40:1089-1096(1996).
 DR EMBL; D87558; BAA21713.1; -;
 DR HSSP; P02468; LTLE.
 DR INTERPRO; IPR000083; -;
 DR INTERPRO; IPR000152; -;
 DR INTERPRO; IPR000561; -;
 DR INTERPRO; IPR000742; -;
 DR INTERPRO; IPR001010; -;
 DR INTERPRO; IPR001438; -;
 DR INTERPRO; IPR001774; -;
 DR INTERPRO; IPR001881; -;
 DR INTERPRO; IPR001947; -;
 DR INTERPRO; IPR002049; -;
 DR PFAM; PF00008; EGF; 14.
 DR PFAM; PF01414; DSL; 1.
 DR PRINTS; PR00010; EGFBLD.
 DR PRINTS; PR00011; EGFAMININ.
 DR PRINTS; PR00012; FNTYPEI.
 DR PRINTS; PR00286; CHARVDTXIN.
 DR PRINTS; PR00287; THLONIN.
 DR PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_10.
 DR PROSITE; PS00022; EGF_1; UNKNOWN_16.
 DR PROSITE; PS01186; EGF_2; 10.
 DR PROSITE; PS01187; EGF_CA; 8.
 KW Glycoprotein; EGF-like domain.
 FT NON_TER
 SQ SEQUENCE 1212 AA; 134188 MW; BC44D29F1C7985FE CRC64;

Query Match 3.6%; Score 8; DB 13; Length 1212;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 80 PNPCHNGG 87
 Db 654 PNPCHNGG 661

RESULT 22
 Q9NS15 PRELIMINARY; PRT; 1242 AA.
 ID Q9NS15
 AC Q9NS15
 DT 01-OCT-2000 (TReMBLrel. 15, Created)
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
 DE LATENT TRANSFORMING GROWTH FACTOR BETA BINDING PROTEIN 3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Saharinen J., Penttinen C., Keski-Oja J.;
 RT "Cloning of human LTBP-3.";
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF135960; AAF62352.2; -;
 SQ SEQUENCE 1242 AA; 133028 MW; 155C87FB69AB221B CRC64;

Query Match 3.6%; Score 8; DB 4; Length 1242;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 121 NECEAEPC 128
 Db 604 NECEAEPC 611

RESULT 23
 Q61810 PRELIMINARY; PRT; 1253 AA.
 ID Q61810
 AC Q61810
 DT 01-NOV-1996 (TReMBLrel. 01, Created)
 DT 01-JUL-1997 (TReMBLrel. 04, Last sequence update)
 DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
 DE LATENT TRANSFORMING GROWTH FACTOR-BETA BINDING PROTEIN.
 GN LTBP3 OR LTBP-3.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95247723; PubMed=7730318;
 RA Yin W., Smiley E., Germiller J., Mecham R.P., Florer J.B.,
 RA Wenstrup R.J., Bonadio J.;
 RT "Isolation of a novel latent transforming growth factor-beta binding
 protein gene (LTBP-3).";
 RL J. Biol. Chem. 270:10147-10160(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Bonadio J.;
 RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; L40459; AAB53015.1; -;
 DR HSSP; P07204; 1ADX.
 DR MGD; MGI:1101355; Ltbp3.
 DR INTERPRO; IPR000152; -;
 DR INTERPRO; IPR000561; -;
 DR INTERPRO; IPR001881; -;
 DR INTERPRO; IPR002212; -;
 DR PFAM; PF00008; EGF; 13.
 DR PFAM; PF00683; TB; 3
 DR PROSITE; PS00010; ASX_HYDROXYL; 10.
 DR PROSITE; PS00022; EGF_1; UNKNOWN_2.
 DR PROSITE; PS01186; EGF_2; 5.
 DR PROSITE; PS01187; EGF_CA; 11.
 KW Glycoprotein; EGF-like domain.
 SQ SEQUENCE 1253 AA; 134705 MW; 3DD4521A3DF5F7A7 CRC64;

Query Match 3.6%; Score 8; DB 11; Length 1253;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 121 NECEAEPC 128
 Db 598 NECEAEPC 605

RESULT 24
 Q07314 PRELIMINARY; PRT; 1378 AA.
 ID Q07314
 AC Q07314
 DT 01-NOV-1996 (TReMBLrel. 01, Created)
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
 DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
 DE SECRETED NEUREXIN III-ALPHA-C PRECURSOR.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

ON NCBI_TaxID=10116;
[1]
RX SEQUENCE FROM N.A.
RC STRAIN-SPRAGUE-DAWLEY; TISSUE-BRAIN;
RX MEDLINE=93342001; PubMed=8341647;
RA Ushkaryov Y.A., Sudhof T.C.;
RT "Neurexin III alpha: extensive alternative splicing generates
RT membrane-bound and soluble forms.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:6410-6414(1993).
CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.
DR EMBL: L14851; AAA02854.1; -;
DR INTERPRO: IPR000152; -;
DR INTERPRO: IPR000561; -;
DR INTERPRO: IPR000991; -;
DR INTERPRO: IPR001791; -;
DR PFAM: PF00008; EGF; 3.
DR PFAM: PF00054; laminin_G; 6.
DR PROSITE: PS00010; ASX_HYDROXYL; UNKNOWN_1.
DR PROSITE: PS00442; GATASE_TYPE_1; UNKNOWN_1.
KW Alternative splicing; signal.
FT SIGNAL 1 27 POTENTIAL.
FT CHAIN 28 1378 SECRETED NEUREXIN III-ALPHA-C.
FT VARIANT 4 4 T -> S.
FT VARIANT 1210 1210 G -> S.
SQ SEQUENCE 1378 AA; 151960 MW; FB88CCC4F7199BCB CRC64;

Query Match 3.6%; Score 8; DB 11; Length 1378;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 PCENGGIC 37
|||||||
Db 206 PCENGGIC 213

RESULT 25
Q07313
ID Q07313 PRELIMINARY; PRT; 1395 AA.
AC Q07313;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE SECRETED NEUREXIN III-ALPHA-B PRECURSOR.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SPRAGUE-DAWLEY; TISSUE-BRAIN;
RX MEDLINE=93342001; PubMed=8341647;
RA Ushkaryov Y.A., Sudhof T.C.;
RT "Neurexin III alpha: extensive alternative splicing generates
RT membrane-bound and soluble forms.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:6410-6414(1993).
CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.
DR EMBL: L14851; AAA02855.1; -;
DR INTERPRO: IPR000152; -;
DR INTERPRO: IPR000561; -;
DR INTERPRO: IPR000991; -;
DR INTERPRO: IPR001791; -;
DR PFAM: PF00008; EGF; 3.
DR PFAM: PF00054; laminin_G; 6.
DR PROSITE: PS00010; ASX_HYDROXYL; UNKNOWN_1.
DR PROSITE: PS00442; GATASE_TYPE_1; UNKNOWN_1.
KW Alternative splicing; signal.
FT SIGNAL 1 27 POTENTIAL.
FT CHAIN 28 1395 SECRETED NEUREXIN III-ALPHA-B.
FT VARIANT 4 4 T -> S.
FT VARIANT 1210 1210 G -> S.
SQ SEQUENCE 1395 AA; 153888 MW; F51D7F30DADE9E49 CRC64;

Query Match 3.6%; Score 8; DB 11; Length 1395;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 PCENGGIC 37
|||||||
Db 206 PCENGGIC 213

Search completed: May 23, 2001, 06:24:43
Job time: 197 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 06:21:18 ; Search time 35.38 seconds
(without alignments)
357.068 Million cell updates/sec

Title: US-09-237-981-29

Perfect score: 221

Sequence: 1 MKHLVAAMLLVGLSLGVPQF.....INAWTAENDRWPIQVTVG 221

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 390729 seqs, 57163235 residues

Word size : 0

Total number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 50 summaries

Database : A_Geneseq_0401.*

1:	/SID86/gcgdata/geneseq/geneseq/AA1980.DAT.*
2:	/SID86/gcgdata/geneseq/geneseq/AA1981.DAT.*
3:	/SID86/gcgdata/geneseq/geneseq/AA1982.DAT.*
4:	/SID86/gcgdata/geneseq/geneseq/AA1983.DAT.*
5:	/SID86/gcgdata/geneseq/geneseq/AA1984.DAT.*
6:	/SID86/gcgdata/geneseq/geneseq/AA1985.DAT.*
7:	/SID86/gcgdata/geneseq/geneseq/AA1986.DAT.*
8:	/SID86/gcgdata/geneseq/geneseq/AA1987.DAT.*
9:	/SID86/gcgdata/geneseq/geneseq/AA1988.DAT.*
10:	/SID86/gcgdata/geneseq/geneseq/AA1989.DAT.*
11:	/SID86/gcgdata/geneseq/geneseq/AA1990.DAT.*
12:	/SID86/gcgdata/geneseq/geneseq/AA1991.DAT.*
13:	/SID86/gcgdata/geneseq/geneseq/AA1992.DAT.*
14:	/SID86/gcgdata/geneseq/geneseq/AA1993.DAT.*
15:	/SID86/gcgdata/geneseq/geneseq/AA1994.DAT.*
16:	/SID86/gcgdata/geneseq/geneseq/AA1995.DAT.*
17:	/SID86/gcgdata/geneseq/geneseq/AA1996.DAT.*
18:	/SID86/gcgdata/geneseq/geneseq/AA1997.DAT.*
19:	/SID86/gcgdata/geneseq/geneseq/AA1998.DAT.*
20:	/SID86/gcgdata/geneseq/geneseq/AA1999.DAT.*
21:	/SID86/gcgdata/geneseq/geneseq/AA2000.DAT.*
22:	/SID86/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	221	100.0	221	20	W94684
2	194	87.8	221	18	W10366
3	194	87.8	480	18	W10364
4	194	87.8	480	20	W94683
5	45	20.4	102	18	W10370
6	45	20.4	481	18	W10365
7	45	20.4	513	20	W94685
8	41	18.6	43	20	W94687
9	18	8.1	25	18	W10368
10	18	8.1	25	20	W94698
11	18	8.1	42	20	W94688

12	18	8.1	321	20	W94697
13	14	6.3	57	20	W94686
14	10	4.5	1010	18	W87896
15	10	4.5	1036	18	W18351
16	10	4.5	1187	18	W18352
17	10	4.5	1193	17	W05835
18	10	4.5	1193	21	W59599
19	10	4.5	1208	19	W40827
20	10	4.5	1218	17	W05833
21	10	4.5	1218	18	W18354
22	10	4.5	1218	19	W44301
23	10	4.5	1218	20	W87894
24	10	4.5	1218	21	W59597
25	9	4.1	1404	14	R38304
26	9	4.1	1404	21	Y59600
27	9	4.1	1530	21	B28151
28	8	3.6	211	21	B38466
29	8	3.6	694	21	B42900
30	8	3.6	833	13	R28960
31	8	3.6	834	21	Y79031
32	8	3.6	908	16	R85443
33	8	3.6	912	16	R85442
34	8	3.6	1208	21	Y70551
35	8	3.6	1251	16	R79475
36	8	3.6	1257	21	Y70554
37	7	3.2	40	21	B38468
38	7	3.2	46	17	R96227
39	7	3.2	46	17	R96224
40	7	3.2	46	17	R96226
41	7	3.2	60	21	B28303
42	7	3.2	60	21	Y86285
43	7	3.2	114	14	R41799
44	7	3.2	114	16	R84184
45	7	3.2	114	16	W01599
46	7	3.2	115	11	R07074
47	7	3.2	115	13	R22004
48	7	3.2	115	17	R94610
49	7	3.2	116	11	R06831
50	7	3.2	118	11	R06833

ALIGNMENTS

RESULT 1
W94684
ID W94684 standard; Protein; 221 AA.
XX
AC W94684;
XX
DT 05-MAY-1999 (first entry)
XX
DE Truncated murine Del-1 protein.
XX
KW Del-1; developmentally-regulated endothelial cell locus 1; cancer;
KW discoidin I; factor VIII-like domain; epidermal growth factor; EGF;
KW diabetic retinopathy; rheumatoid arthritis; endometriosis;
KW angiogenesis.
XX
OS Mus sp.
XX
FN US5877281-A.
XX
PD 02-MAR-1999.
XX
PF 05-JUN-1996; 96US-0659235.
XX
PR 05-JUN-1996; 96US-0659235.
XX
PR 07-JUN-1995; 95US-0480229.
XX
PA (PROG-) PROGENITOR INC.
XX (UYVA-) UNIV VANDERBILT.

PI Hogan B, Quertermous T, Snodgrass HR, Zupancic TJ;
 XX WPI; 1999-189720/16.
 DR N-PSDB; X18507.
 XX
 XX Del-1 polypeptide sequences - useful for treatment of cancer,
 PT diabetic retinopathy, rheumatoid arthritis and endometriosis
 XX
 XX Claim 3; Column 71-72; 73pp; English.
 PS
 XX The present sequence is truncated murine developmentally-regulated
 CC endothelial cell locus 1 (del-1). The protein has epidermal growth factor
 CC (EGF) like domains and discoidin I/factor VIII-like domains. The Del-1
 CC proteins have an inhibitory effect on angiogenesis (blood vessel growth),
 CC this activity may be useful clinically to prevent neovascularisation of
 CC tissues such as tumour nodules and prevention of metastases. The anti-
 CC angiogenic activity of Del-1 may be used to treat abnormal conditions
 CC that result from angiogenesis, including cancer, diabetic retinopathy,
 CC rheumatoid arthritis and endometriosis. Since Del-1 promotes
 CC angiogenesis it can be used to treat cardiac ischaemia, thrombotic
 CC stroke, wound healing and peripheral vascular disease. Del-1 is also
 CC useful for promoting bone formation. Del-1 binds to alpha V beta 3
 CC integrin, and is an apoptosis inducer.
 XX
 XX Sequence 221 AA;

Query Match 100.0%; Score 221; DB 20; Length 221;
 Best Local Similarity 100.0%; Pred. No. 1.1e-202; Indels 0; Gaps 0;
 Matches 221; Conservative 0; Mismatches 0;
 QY 1 MKHLVAALLVGLSLGVPOFGKGDICNPENPCENGICLSGLADDSFCECFAGPNC 60
 DB 1 mkhlvaallvlgslgvpo fgk g d i c n p e n p c e n g i c l s g l a d d s f c e c p e f a g p n c s 60
 QY 61 SVVEVASDEEKPTSAGPCIPNCHNGTCEISEAYRGDTFFIGYVCKPGRFGNGIHCHQNI 120
 DB 61 svvevasdeekptsagpcipnchngtceiseayrgdtffigyvckpgrfngihcqnhl 120
 QY 121 NECEAPCRNGGICTDLVANYSCPCPEFMGRNCQKCSGHLGIEGGIISNQOITASSNH 180
 DB 121 neceapcrnggictdlvanyscpcpefmgrncqkcsghlgieggiiisnqqitassnh 180
 QY 181 RALFGLQKWYPYAAALNKKGLINAWTAENDRPWQIVTVG 221
 DB 181 ralfglqkwypyaaalnkkglinawtaendrpwqivtvvg 221

RESULT 2

W10366
 ID W10366 standard; Protein; 221 AA.

AC W10366;

XX 03-MAY-1997 (first entry)

XX Murine Del-1 truncated minor protein.

DE Del-1; developmentally-regulated endothelial cell locus-1;
 KW signal transduction; cancer; tumour marker; angiogenesis;
 KW diabetic retinopathy; rheumatoid arthritis; endometriosis;
 KW cardiac ischaemia; stroke; vascular disease; wound healing;
 KW vulnery; bone formation; diagnosis; therapy.

OS Mus sp.

XX Key Location/Qualifiers
 PH Peptide 1..23
 FT Peptide /label= Sig_peptide
 FT Protein 24..480
 FT Mat_protein /label= Mat_protein
 FT Domain 26..59
 FT /label= EGF-1

FT Domain /note= "epidermal growth factor-like domain 1"
 FT 78..116
 FT /label= EGF-2
 FT /note= "epidermal growth factor-like domain 2"
 FT 123..154
 FT /label= EGF-3
 FT /note= "epidermal growth factor-like domain 3"
 FT 158..221
 FT /label= Discoidin-1
 FT /note= "truncated discoidin I/factor VIII-like
 FT domain 1"
 XX
 XX W09640769-A1.
 PN
 XX 19-DEC-1996.
 PD
 XX 05-JUN-1996; 96WO-US09456.
 PF
 XX 07-JUN-1995; 95US-0480229.
 PR
 XX (PROG-) PROGENITOR INC.
 PA (UYVA-) UNIV VANDERBILT.
 XX
 XX Hogan B, Quertermous T, Snodgrass HR, Zupancic TJ;
 PI WPI; 1997-052233/05;
 XX N-PSDB; T47339.
 DR
 XX New developmentally regulated endothelial cell locus-1 (del-1) gene
 PT - used to develop prods. for the diagnosis and treatment of cancer
 PT and conditions involving abnormal angiogenesis
 XX
 XX Example; Fig 12; 137pp; English.
 PS
 XX A truncated version (W10366) of murine developmentally-regulated
 CC endothelial cell locus-1 (del-1) gene product (see also W10364)
 CC contains a signal peptide, all 3 EGF-like domains but only a
 CC partial N-terminal discoidin I/factor VIII-like domain (about 40%).
 CC It is the product of a murine del-1 minor sequence (T47339). This
 CC transcript was cloned only from mouse embryonic libraries, but was
 CC verified through cloning of several independent cDNAs.
 XX
 XX Sequence 221 AA;

Query Match 87.8%; Score 194; DB 18; Length 221;
 Best Local Similarity 100.0%; Pred. No. 5.5e-177;
 Matches 194; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKHLVAALLVGLSLGVPOFGKGDICNPENPCENGICLSGLADDSFCECFAGPNC 60
 DB 1 mkhlvaallvlgslgvpo fgk g d i c n p e n p c e n g i c l s g l a d d s f c e c p e f a g p n c s 60
 QY 61 SVVEVASDEEKPTSAGPCIPNCHNGTCEISEAYRGDTFFIGYVCKPGRFGNGIHCHQNI 120
 DB 61 svvevasdeekptsagpcipnchngtceiseayrgdtffigyvckpgrfngihcqnhl 120
 QY 121 NECEAPCRNGGICTDLVANYSCPCPEFMGRNCQKCSGHLGIEGGIISNQOITASSNH 180
 DB 121 neceapcrnggictdlvanyscpcpefmgrncqkcsghlgieggiiisnqqitassnh 180
 QY 181 RALFGLQKWYPYAA 194
 DB 181 ralfglqkwypyya 194

RESULT 3

W10364
 ID W10364 standard; Protein; 480 AA.
 XX
 AC W10364;
 XX
 XX 03-MAY-1997 (first entry)
 DT


```

XX Mouse developmentally-regulated endothelial cell locus-1 protein.
DE Del-1; developmentally-regulated endothelial cell locus-1;
KW signal transduction; cancer; tumour marker; angiogenesis;
KW diabetic retinopathy; rheumatoid arthritis; endometriosis;
KW cardiac ischaemia; stroke; vascular disease; wound healing;
KW vulnary; bone formation; diagnosis; therapy.
OS
XX
XX Mus sp.
FH Key Location/Qualifiers
FT Peptide 1..23
FT Protein /label= Sig_peptide
FT /label= Mat_protein
FT Domain 26..59
FT /label= EGF-1
FT /note= "epidermal growth factor-like domain 1"
FT Domain 78..116
FT /label= EGF-2
FT /note= "epidermal growth factor-like domain 2"
FT Domain 123..154
FT /label= EGF-3
FT /note= "epidermal growth factor-like domain 3"
FT Domain 158..314
FT /label= Discoidin-1
FT /note= "discoidin I/factor VIII-like domain 1"
FT Domain 319..476
FT /label= Discoidin-2
FT /note= "discoidin I/factor VIII-like domain 2"
XX
XX W09640769-A1.
XX
XX 19-DEC-1996.
XX
XX 05-JUN-1996; 96WO-US09456.
XX
XX 07-JUN-1995; 95US-0480229.
XX
XX (PROG-) PROGENITOR INC.
XX (UYVA-) UNIV VANDERBILT.
XX
XX Hogan B, Quatermous T, Snodgrass HR, Zupancic TJ;
XX WPI; 1997-052233/05.
XX N-PSDB; T47338.
XX
XX New developmentally regulated endothelial cell locus-1 (del-1) gene
XX - used to develop prods. for the diagnosis and treatment of cancer
XX and conditions involving abnormal angiogenesis
XX
XX Claim 3; Fig 6; 137pp; English.
XX
XX Murine Del-1 (W10364) is the polypeptide product of the murine
XX developmentally-regulated endothelial cell locus-1 (del-1) gene
XX (T47338). It shows 94% amino acid homology to the human Del-1
XX protein (W10365). Structurally, members of this novel gene family
XX contain 3 EGF-like domains and 2 discoidin I/factor VIII-like
XX domains. Del-1 is expressed in endothelial and certain tumour
XX cells. Its ability to inhibit vascular formation allows its used
XX as an anti-angiogenic agent. It can be used as a tumour marker,
XX to identify Del-1 binding partners, and to modulate endothelial
XX cell growth and blood vessel formation. Recombinant Del-1 can be
XX produced in transformed host cells utilising vectors incorporating
XX del-1 nucleic acids.
XX
XX Sequence 480 AA;
XX
XX Query Match 87.8%; Score 194; DB 18; Length 480;
XX Best Local Similarity 100.0%; Pred. No. 1.1e-176;
XX Matches 194; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MKHLVAALLVGLGVPOFGKGDICNPENPCENGICLSGLADDSFSCPCPGFAGPNC 60
DB 1 mkhlvaallvglslgvpqfgkdgicnnpncengiclsgladddsfscpcpgfagpncs 60
QY 61 SVVEVASDEEKPTTSAGPCIPNPNCHNGGTCEISEAYRGDTFIGYVCKPRGFIHQHNI 120
DB 61 svvevasdeekptsagpcipnpchnggtceiseayrgdtfigyvcckprgfighcnhni 120
QY 121 NECEAEPCRNNGICTDLVANYSCPCPGFMGRNCQKCSGHLGIEGGIISNQOITASSNH 180
DB 121 neceaeprnngictdlvanyscpcepfmgrncqkcsghlgieggiiisnqgitassnh 180
QY 181 RALFLGLQKWYPYA 194
DB 181 ralflglqkwypya 194
RESULT 4
W94683
ID W94683 standard; Protein; 480 AA.
XX
AC W94683;
XX
XX 05-MAY-1999 (first entry)
DE Murine Del-1 protein.
XX
XX Del-1; developmentally-regulated endothelial cell locus 1; cancer;
KW discoidin 1; factor VIII-like domain; epidermal growth factor; EGF;
KW diabetic retinopathy; rheumatoid arthritis; endometriosis;
XX
XX Mus sp.
XX
XX US5877281-A.
XX
XX 02-MAR-1999.
XX
XX 05-JUN-1996; 96US-0659235.
XX
XX 05-JUN-1996; 96US-0659235.
XX 07-JUN-1995; 95US-0480229.
XX
XX (PROG-) PROGENITOR INC.
XX (UYVA-) UNIV VANDERBILT.
XX
XX Hogan B, Quatermous T, Snodgrass HR, Zupancic TJ;
XX WPI; 1999-189720/16.
XX N-PSDB; X18506.
XX
XX Del-1 polypeptide sequences - useful for treatment of cancer,
XX diabetic retinopathy, rheumatoid arthritis and endometriosis
XX
XX Claim 1; Fig 3; 73pp; English.
XX
XX The present sequence is murine developmentally-regulated endothelial
XX cell locus 1 (del-1). The protein has epidermal growth factor (EGF) like
XX domains and discoidin I/factor VIII-like domains. The del-1 proteins
XX have an inhibitory effect on angiogenesis (blood vessel growth), this
XX activity may be useful clinically to prevent neovascularisation of
XX tissues such as tumour nodules and prevention of metastases. The anti-
XX angiogenic activity of del-1 may be used to treat abnormal conditions
XX that result from angiogenesis, including cancer, diabetic retinopathy,
XX rheumatoid arthritis and endometriosis. Since Del-1 promotes
XX angiogenesis it can be used to treat cardiac ischaemia, thrombotic
XX stroke, wound healing and peripheral vascular disease. Del-1 is also
XX useful for promoting bone formation. Del-1 binds to alpha V beta 3
XX integrin, and is an apoptosis inducer.
XX
XX Sequence 480 AA;

```


XX New developmentally regulated endothelial cell locus-1 (del-1) gene
PT - used to develop prods. for the diagnosis and treatment of cancer
PT and conditions involving abnormal angiogenesis
XX
PS Claim 4; Fig 6; 137pp; English.
XX
CC Human Del-1 (W10365) is the polypeptide product of the human
CC developmentally-regulated endothelial cell locus-1 (del-1) gene
CC (T47343). It shows 94% amino acid homology to the mouse Del-1
CC protein (W10364). Structurally, members of this novel gene family
CC contain 3 EGF-like domains and 2 discoidin I/factor VIII-like
CC domains. Del-1 is expressed in endothelial and certain tumour
CC cells. Its ability to inhibit vascular formation allows its used
CC as an anti-angiogenic agent. It can be used as a tumour marker,
CC to identify Del-1 binding partners, and to modulate endothelial
CC cell growth and blood vessel formation. Recombinant Del-1 can be
CC produced in transformed host cells utilising vectors incorporating
CC del-1 nucleic acids.
XX
SQ Sequence 481 AA;

Query Match 20.4%; Score 45; DB 18; Length 481;
Best Local Similarity 100.0%; Pred. No. 8.7e-35;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 PNPCHNGTCEISEAYRGDTFYGVCPCPRGFNGHCHQHINECE 124
|||||
DB 80 pnpchnggtceiseayrgdtfgyvckprgfgnhcqhninece 124

RESULT 7
W94685
ID W94685 standard; Protein; 513 AA.
XX
AC W94685;
XX
DT 05-MAY-1999 (first entry)
XX
DE Human Del-1 protein.
XX
KW Del-1; developmentally-regulated endothelial cell locus 1; cancer;
KW discoidin I; factor VIII-like domain; epidermal growth factor; EGF;
KW diabetic retinopathy; rheumatoid arthritis; endometriosis;
KW angiogenesis.
XX
OS Homo sapiens.
XX
PN US5877281-A.
XX
PD 02-MAR-1999.
XX
PF 05-JUN-1996; 96US-0659235.
XX
PR 05-JUN-1996; 96US-0659235.
XX
PR 07-JUN-1995; 95US-0480229.
XX
PA (PROG-) PROGENITOR INC.
PA (UYVA-) UNIV VANDERBILT.
XX
PI Hogan B, Quertermous T, Snodgrass HR, Zupancic TJ;
XX
DR WPI; 1999-189720/16.
DR N-PSDB; X18508.
XX
PT Del-1 polypeptide sequences - useful for treatment of cancer,
PT diabetic retinopathy, rheumatoid arthritis and endometriosis
XX
PS Claim 5; Fig 4; 73pp; English.
XX
XX The present sequence is human developmentally-regulated endothelial
CC cell locus 1 (Del-1). The protein has epidermal growth factor (EGF) like

CC domains and discoidin I/factor VIII-like domains. The Del-1 proteins
CC have an inhibitory effect on angiogenesis (blood vessel growth), this
CC activity may be useful clinically to prevent neovascularisation of
CC tissues such as tumour nodules and prevention of metastases. The anti-
CC angiogenic activity of Del-1 may be used to treat abnormal conditions
CC that result from angiogenesis, including cancer, diabetic retinopathy,
CC rheumatoid arthritis and endometriosis. Since Del-1 promotes
CC angiogenesis it can be used to treat cardiac ischaemia, thrombotic
CC stroke, wound healing and peripheral vascular disease. Del-1 is also
CC useful for promoting bone formation. Del-1 binds to alpha V beta 3
CC integrin, and is an apoptosis inducer.
XX
SQ Sequence 513 AA;

Query Match 20.4%; Score 45; DB 20; Length 513;
Best Local Similarity 100.0%; Pred. No. 9.2e-35;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 PNPCHNGTCEISEAYRGDTFYGVCPCPRGFNGHCHQHINECE 124
|||||
DB 112 pnpchnggtceiseayrgdtfgyvckprgfgnhcqhninece 156

RESULT 8
W94687
ID W94687 standard; peptide; 43 AA.
XX
AC W94687;
XX
DT 05-MAY-1999 (first entry)
XX
DE Del-1 epidermal growth factor like domain #2.
XX
KW Del-1; developmentally-regulated endothelial cell locus 1; cancer;
KW discoidin I; factor VIII-like domain; epidermal growth factor; EGF;
KW diabetic retinopathy; rheumatoid arthritis; endometriosis;
KW angiogenesis.
XX
OS Mus sp.
OS Synthetic.
XX
PN US5877281-A.
XX
PD 02-MAR-1999.
XX
PF 05-JUN-1996; 96US-0659235.
XX
PR 05-JUN-1996; 96US-0659235.
XX
PR 07-JUN-1995; 95US-0480229.
XX
PA (PROG-) PROGENITOR INC.
PA (UYVA-) UNIV VANDERBILT.
XX
PI Hogan B, Quertermous T, Snodgrass HR, Zupancic TJ;
XX
DR WPI; 1999-189720/16.
XX
PT Del-1 polypeptide sequences - useful for treatment of cancer,
PT diabetic retinopathy, rheumatoid arthritis and endometriosis
XX
PS Claim 40; Fig 10; 73pp; English.
XX
XX The present sequence represents an epidermal growth factor like domain
CC from developmentally-regulated endothelial cell locus 1 (Del-1). Del-1
CC also has discoidin I/factor VIII-like domains. The Del-1 proteins
CC have an inhibitory effect on angiogenesis (blood vessel growth), this
CC activity may be useful clinically to prevent neovascularisation of
CC tissues such as tumour nodules and prevention of metastases. The anti-
CC angiogenic activity of Del-1 may be used to treat abnormal conditions
CC that result from angiogenesis, including cancer, diabetic retinopathy,
CC rheumatoid arthritis and endometriosis. Since Del-1 promotes
CC angiogenesis it can be used to treat cardiac ischaemia, thrombotic


```

RESULT 11
W94688
ID W94688 standard; peptide; 42 AA.
XX
AC W94688;
XX
DT 05-MAY-1999 (first entry)
XX
DE Del-1 epidermal growth factor like domain #3.
XX
KW Del-1; developmentally-regulated endothelial cell locus 1; cancer;
KW discoidin I; factor VIII-like domain; epidermal growth factor; EGF;
KW diabetic retinopathy; rheumatoid arthritis; endometriosis;
KW angiogenesis.
XX
OS Mus sp.
OS Synthetic.
XX
PN US5877281-A.
XX
PD 02-MAR-1999.
XX
PF 05-JUN-1996; 96US-0659235.
XX
PR 05-JUN-1996; 96US-0659235.
PR 07-JUN-1995; 95US-0480229.
XX
PA (PROG-) PROGENITOR INC.
PA (UYVA-) UNIV VANDERBILT.
XX
Hogan B, Quertermous T, Snodgrass HR, Zupancic TJ;
WPI; 1999-189720/16.
XX
Del-1 polypeptide sequences - useful for treatment of cancer,
PT diabetic retinopathy, rheumatoid arthritis and endometriosis
XX
Claim 40; Fig 10; 73pp; English.
XX
The present sequence represents an epidermal growth factor like domain
CC from developmentally-regulated endothelial cell locus 1 (Del-1). Del-1
CC also has discoidin I/factor VIII-like domains. The Del-1 proteins
CC have an inhibitory effect on angiogenesis (blood vessel growth), this
CC activity may be useful clinically to prevent neovascularisation of
CC tissues such as tumour nodules and prevention of metastases. The anti-
CC angiogenic activity of Del-1 may be used to treat abnormal conditions
CC that result from angiogenesis, including cancer, diabetic retinopathy,
CC rheumatoid arthritis and endometriosis. Since Del-1 promotes
CC angiogenesis it can be used to treat cardiac ischaemia, thrombotic
CC stroke, wound healing and peripheral vascular disease. Del-1 is also
CC useful for promoting bone formation. Del-1 binds to alpha V beta 3
CC integrin, and is an apoptosis inducer.
XX
SQ Sequence 42 AA;

Query Match 8.1%; Score 18; DB 20; Length 42;
Best Local Similarity 100.0%; Pred. No. 5.2e-10;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 140 NYSCECPGEGFMGRNCQYK 157
|||||
Db 25 nyscecpgefmgncqyk 42

RESULT 12
W94697
ID W94697 standard; Protein; 321 AA.
XX
AC W94697;
XX
DT 05-MAY-1999 (first entry)
XX

```

```

DE Human milk fat globule protein MFG-E8.
XX
KW Del-1; developmentally-regulated endothelial cell locus 1; cancer;
KW discoidin I; factor VIII-like domain; epidermal growth factor; EGF;
KW diabetic retinopathy; rheumatoid arthritis; endometriosis;
KW angiogenesis.
XX
OS Homo sapiens.
XX
PN US5877281-A.
XX
PD 02-MAR-1999.
XX
PF 05-JUN-1996; 96US-0659235.
XX
PR 05-JUN-1996; 96US-0659235.
PR 07-JUN-1995; 95US-0480229.
XX
PA (PROG-) PROGENITOR INC.
PA (UYVA-) UNIV VANDERBILT.
XX
Hogan B, Quertermous T, Snodgrass HR, Zupancic TJ;
WPI; 1999-189720/16.
XX
Del-1 polypeptide sequences - useful for treatment of cancer,
PT diabetic retinopathy, rheumatoid arthritis and endometriosis
XX
Example; Column 63-64; 73pp; English.
XX
The present sequence represents human milk fat globule protein MFG-E8,
CC which has homology to the developmentally-regulated endothelial cell
CC locus 1 (Del-1). The Del-1 protein has epidermal growth factor (EGF) like
CC domains and discoidin I/factor VIII-like domains. The Del-1 proteins
CC have an inhibitory effect on angiogenesis (blood vessel growth), this
CC activity may be useful clinically to prevent neovascularisation of
CC tissues such as tumour nodules and prevention of metastases. The anti-
CC angiogenic activity of Del-1 may be used to treat abnormal conditions
CC that result from angiogenesis, including cancer, diabetic retinopathy,
CC rheumatoid arthritis and endometriosis. Since Del-1 promotes
CC angiogenesis it can be used to treat cardiac ischaemia, thrombotic
CC stroke, wound healing and peripheral vascular disease. Del-1 is also
CC useful for promoting bone formation. Del-1 binds to alpha V beta 3
CC integrin, and is an apoptosis inducer.
XX
SQ Sequence 321 AA;

Query Match 8.1%; Score 18; DB 20; Length 321;
Best Local Similarity 100.0%; Pred. No. 3.1e-09;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 196 LNKKGILINAWTAENDRW 213
|||||
Db 39 lnnkglinawtaendrw 56

RESULT 13
W94686
ID W94686 standard; peptide; 57 AA.
XX
AC W94686;
XX
DT 05-MAY-1999 (first entry)
XX
DE Del-1 epidermal growth factor like domain #1.
XX
KW Del-1; developmentally-regulated endothelial cell locus 1; cancer;
KW discoidin I; factor VIII-like domain; epidermal growth factor; EGF;
KW diabetic retinopathy; rheumatoid arthritis; endometriosis;
KW angiogenesis.
XX
OS Mus sp.

```

```
OS Synthetic.
XX US5877281-A.
XX PD 02-MAR-1999.
XX PF 05-JUN-1996; 96US-0659235.
XX PR 05-JUN-1996; 96US-0659235.
XX PR 07-JUN-1995; 95US-0480229.
XX (PROG-) PROGENITOR INC.
XX (UYVA-) UNIV VANDERBILT.
XX Hogan B, Quertermous T, Snodgrass HR, Zupancic TJ;
XX WPI; 1999-189720/16.
XX DR Del-1 polypeptide sequences - useful for treatment of cancer,
XX PT diabetic retinopathy, rheumatoid arthritis and endometriosis
XX PS Claim 40; Fig 10; 73pp; English.
XX
XX The present sequence represents an epidermal growth factor like domain
XX from developmentally-regulated endothelial cell locus 1 (Del-1). Del-1
XX also has discoidin I/factor VIII-like domains. The Del-1 proteins
XX have an inhibitory effect on angiogenesis (blood vessel growth); this
XX activity may be useful clinically to prevent neovascularisation of
XX tissues such as tumour nodules and prevention of metastases. The anti-
XX angiogenic activity of Del-1 may be used to treat abnormal conditions
XX that result from angiogenesis, including cancer, diabetic retinopathy,
XX rheumatoid arthritis and endometriosis. Since Del-1 promotes
XX angiogenesis it can be used to treat cardiac ischaemia, thrombotic
XX stroke, wound healing and peripheral vascular disease. Del-1 is also
XX useful for promoting bone formation. Del-1 binds to alpha v beta 3
XX integrin, and is an apoptosis inducer.
XX
XX Sequence 57 AA:
XX
XX Query Match 6.3%; Score 14; DB 20; Length 57;
XX Best Local Similarity 100.0%; Pred. No. 4.4e-06;
XX Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 57 PNCSSVVEVASDEE 70
XX |||||
XX Db 37 pncssvvevasdee 50
XX
XX RESULT 14
XX W87896
XX ID W87896 standard; Peptide; 1010 AA.
XX AC W87896;
XX XX 26-APR-1999 (first entry)
XX DE Human JAGGED1 soluble polypeptide.
XX KW JAGGED1; JAGGED1; human; notch ligand; stem cell;
XX KW progenitor cell; haematopoiesis; cell differentiation;
XX KW Alagille syndrome; leukaemia; lymphoma; diagnosis; therapy.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT Peptide 1..21
XX FT /note= "signal peptide"
XX FT Protein 22..1010
XX FT /note= "mature protein"
XX FT Domain 185..239
XX FT /note= "Delta/Serrate/Lag-2 (DSL) domain"
XX FT Region 234..862
XX
XX /note= "EGF-like repeat region"
XX WO9858958-A2.
XX PD 30-DEC-1998.
XX PF 25-JUN-1998; 98WO-US13207.
XX PR 25-JUN-1997; 97US-0882046.
XX (CHIL-) CHILDRENS HOSPITAL PHILADELPHIA.
XX (UNIW) UNIV WASHINGTON.
XX Hood L, Krantz ID, Li L, Spinner NB;
XX WPI; 1999-081220/07.
XX DR New Jagged peptides for inhibiting differentiation of progenitor
XX PT cells - also used for maintaining these cells in undifferentiated
XX PT state, e.g. for haematopoietic reconstitution
XX PS Claim 8; Page 7; 101pp; English.
XX
XX This is the amino acid sequence of a biologically active soluble
XX human JAGGED1 (hJAGGED1) polypeptide comprising amino acid residues
XX 1-1010 of hJAGGED1 (see W87894). It was prepared by PCR
XX amplification (see V63759-60) of hJAGGED1 cDNA (see V63753) and
XX expression in CHO and BHK cells. hJAGGED1 is an activating ligand
XX for Notch protein that is expressed in bone marrow stromal cells.
XX A stromal cell line expressing hJAGGED1 permits survival and
XX proliferation of haematopoietic progenitor cells expressing Notch
XX but inhibits granulocyte differentiation. hJAGGED1 and active
XX peptides can be used (i) to inhibit differentiation of haematopoietic
XX progenitor cells (HPC), e.g. for subsequent production of blood cells
XX for transplantation or dendritic cells for immunotherapy, and (ii) to
XX maintain HPC in the undifferentiated state, particularly totipotent
XX cells or cells able to reconstitute the haematopoietic system, e.g.
XX in patients with leukaemia or lymphoma. Treated HPC, e.g. where
XX taken from a neonate, may be cryopreserved for many years, then
XX thawed for further expansion and differentiation. Optionally
XX JAGGED polypeptides are provided by transformed host cells.
XX (N.B. the amino acid sequence of the polypeptide of W87896 was
XX constructed from the full-length hJAGGED1 amino acid sequence given
XX in Fig1B of the specification).
XX
XX Sequence 1010 AA:
XX
XX Query Match 4.5%; Score 10; DB 20; Length 1010;
XX Best Local Similarity 100.0%; Pred. No. 0.36;
XX Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 80 PNPCHNGGTC 89
XX |||||
XX Db 750 pnpchnggtc 759
XX
XX RESULT 15
XX W18351
XX ID W18351 standard; protein; 1036 AA.
XX AC W18351;
XX XX 11-FEB-1998 (first entry)
XX DE Proliferation and differentiation suppression polypeptide.
XX KW Proliferation; differentiation; suppression; human; delta-1;
XX KW serrate-1; blood cell; neuron; leukaemia; malignant tumour;
XX KW immunosuppression.
XX OS Homo sapiens.
XX XX
```

PN W09719172-A1.

PD 29-MAY-1997.

PF 15-NOV-1996; 96WO-JP03356.

PR 30-NOV-1995; 95JP-0311811.

PR 17-NOV-1995; 95JP-0299611.

XX (ASAH) ASahi KASEI KOGYO KK.

XX Itoh A, Sakano S;

PI WPI; 1997-298110/27.

DR WPI; 1997-298110/27.

XX Peptide(s) encoded by human genes delta-1 and serrate-1 - suppress

PT proliferation and differentiation of undifferentiated human blood

PT cells

XX Claim 5; Page 66-71; 114pp; Japanese.

PS The present sequence represents a polypeptide which suppresses

CC proliferation and differentiation of undifferentiated cells such

CC as neurons and blood cells. The polypeptide may be used for the

CC prevention and control of disorders involving undifferentiated

CC cells, such as leukaemia and malignant tumours, and improvement of

CC blood formation, e.g. after immunosuppression.

XX Sequence 1036 AA;

SQ

Query Match 4.5%; Score 10; DB 18; Length 1036;

Best Local Similarity 100.0%; Pred. No. 0.37;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 80 PNPCHNGGTC 89

|||||

Db 719 pnpchggtc 728

RESULT 16

ID W18352 standard; protein; 1187 AA.

AC W18352;

XX 11-FEB-1998 (first entry)

DT Proliferation and differentiation suppression polypeptide.

XX Proliferation; differentiation; suppression; human; delta-1;

DE serrate-1; blood cell; neuron; leukaemia; malignant tumour;

DE immunosuppression.

KW Homo sapiens.

KW W09719172-A1.

XX 29-MAY-1997.

XX 15-NOV-1996; 96WO-JP03356.

XX 30-NOV-1995; 95JP-0311811.

PR 17-NOV-1995; 95JP-0299611.

XX (ASAH) ASahi KASEI KOGYO KK.

XX Itoh A, Sakano S;

PI WPI; 1997-298110/27.

DR Peptide(s) encoded by human genes delta-1 and serrate-1 - suppress

XX proliferation and differentiation of undifferentiated human blood

PT

PT cells

XX Claim 6; Page 71-76; 114pp; Japanese.

PS The present sequence represents a polypeptide which suppresses

CC proliferation and differentiation of undifferentiated cells such

CC as neurons and blood cells. The polypeptide may be used for the

CC prevention and control of disorders involving undifferentiated

CC cells, such as leukaemia and malignant tumours, and improvement of

CC blood formation, e.g. after immunosuppression.

XX Sequence 1187 AA;

SQ

Query Match 4.5%; Score 10; DB 18; Length 1187;

Best Local Similarity 100.0%; Pred. No. 0.42;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 80 PNPCHNGGTC 89

|||||

Db 719 pnpchggtc 728

RESULT 17

W05835

ID W05835 standard; Protein; 1193 AA.

XX W05835;

XX 28-JAN-1997 (first entry)

DT Chick Serrate.

XX C-Serrate; Notch; cell differentiation; cell fate; tissue repair;

XX central nervous system; cancer; therapy; diagnosis.

XX Gallus sp.

XX Key

XX Location/Qualifiers

XX 1..1041

XX Domain

XX /label= Extracellular_domain

XX 1..5

XX Peptide

XX /label= Sig_peptide

XX /note= "lacks the N-terminal portion owing to

XX truncation of the encoding cDNA clone"

XX Domain

XX 158..203

XX /label= DSL

XX /note= "region of homology with Drosophila Delta

XX and Serrate, predicted to mediate binding

XX with Notch"

XX Domain

XX 208..837

XX /label= ELR

XX /note= "epidermal growth factor-like repeat domain"

XX 208..238

XX /label= ELR1

XX 239..274

XX /label= ELR2

XX 275..313

XX /label= ELR3

XX 314..351

XX /label= ELR4

XX 352..390

XX /label= ELR5

XX 391..427

XX /label= ELR6

XX 428..464

XX /label= ELR7

XX 465..502

XX /label= ELR8

XX 503..540

XX /label= ELR9

XX 541..606

XX /label= ELR10

FT Region 607..644
 FT /label= ELR11
 FT Region 655..682
 FT /label= ELR12
 FT Region 683..721
 FT /label= ELR13
 FT Region 722..759
 FT /label= ELR14
 FT Region 760..797
 FT /label= ELR15
 FT Region 798..837
 FT /label= ELR16
 FT Region 854..911
 FT /label= Cysteine-rich_region
 FT Domain 1042..1066
 FT /label= Transmembrane_domain
 FT Domain 1067..1193
 FT /label= Intracellular_domain
 XX
 PN WO9627610-A1.
 XX
 XX
 PD 12-SEP-1996.
 XX
 XX
 PF 07-MAR-1996; 96WO-US03172.
 XX
 XX
 PR 07-MAR-1995; 95US-0400159.
 XX
 XX (IMCR) IMPERIAL CANCER RES TECHNOLOGY.
 PA (UYIA) UNIV YALE.
 XX
 PI Artavanis-Tsakonas S, Gray GE, Henrique DMP, Ish-Horowitz D;
 PI Lewis JH, Mann RS, Myat AM;
 XX
 XX WPI: 1996-425379/42.
 DR N-PSDB; T40092.
 DR
 XX
 XX Vertebrate Serrate protein and related DNA - used to treat or
 PT prevent malignancies characterised by increased Notch activity.
 PT
 XX Disclosure; Page 112-115; 161pp; English.
 PS
 XX Chicken Serrate (W05835), or C-Serrate, is a ligand for the zygotic
 CC neurogenic locus Notch and is believed to play a major role in
 CC determining cell fates in the central nervous system. Its amino
 CC acid sequence was deduced from a cDNA clone (T40092) obt'd. from an
 CC optic explant cDNA library. C-Serrate is expressed in the central
 CC nervous system, cranial placodes, nephric mesoderm, vascular
 CC system, and limb bud mesenchyme.
 XX
 SQ Sequence 1193 AA;

Query Match 4.5%; Score 10; DB 17; Length 1193;
 Best Local Similarity 100.0%; Pred. No. 0.42;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 80 PNPCHNGGTC 89
 Db 724 pnpchnggtc 733
 |||||

RESULT 18
 Y59599
 ID Y59599 standard; Protein: 1193 AA.
 XX
 AC Y59599;
 XX
 XX 05-APR-2000 (first entry)
 DT
 XX Chick Serrate protein sequence.
 DE
 XX Serrate; cancerous condition; melanoma; lung cancer; breast cancer;
 KW nervous system disorder; infection; nutritional disease; therapy;
 KW

KW cell proliferation promoter; tissue regeneration; chicken.
 XX
 XX Gallus sp.
 XX
 PN US6004924-A.
 XX
 PD 21-DEC-1999.
 XX
 XX 06-MAR-1996; 96US-0611729.
 XX
 XX 11-DEC-1991; 91US-0808458.
 PR 14-SEP-1993; 93US-0121979.
 PR 07-JUN-1994; 94US-0255102.
 PR 07-MAR-1995; 95US-0400159.
 XX
 XX (IMCR) IMPERIAL CANCER RES TECHNOLOGY.
 PA (UYIA) UNIV YALE.
 XX
 PI Ish-Horowitz D, Henrique DMP, Myat AM, Fleming RJ;
 PI Artavanis-Tsakonas S, Gray GE, Mann RS, Lewis JH;
 XX
 DR WPI: 2000-105089/09.
 DR N-PSDB; Z49098.
 XX
 XX Purified Serrate proteins useful for treating neoplasias, nervous
 PT disorders and for promoting cell proliferation and tissue regeneration
 PT and repair -
 PT
 XX Claim 1; Fig 12; 114pp; English.
 PS
 XX This sequence represents a chick serrate protein.
 CC The invention relates to purified vertebrate (mouse, chick, and human)
 CC serrate proteins. The Serrate proteins may be administered to treat a
 CC cancerous condition (e.g. melanoma, lung cancer and breast cancer) by
 CC preventing progression from a preneoplastic or nonmalignant state into a
 CC neoplastic or malignant state. It may also be used to treat nervous
 CC system disorders (such as lesions caused by infections, nutritional
 CC disease and toxic substances) and to promote cell proliferation and
 CC tissue regeneration and repair. The protein itself is administered to
 CC supplement a patient's own production of Serrate proteins (if levels of
 CC expression are low) or to compensate for expression of inactive proteins
 CC due to genetic mutations. The protein may also be used in the production
 CC of antibodies against Serrate proteins which may be used to either down
 CC regulate Serrate activity or to detect Serrate proteins in samples (for
 CC example via enzyme-linked immunosorbent assay (ELISA)). The proteins may
 CC also be used to study Serrate expression and its role in metabolism and
 CC to assay for agents which modulate its expression and activity.
 XX
 SQ Sequence 1193 AA;

Query Match 4.5%; Score 10; DB 21; Length 1193;
 Best Local Similarity 100.0%; Pred. No. 0.42;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 80 PNPCHNGGTC 89
 Db 724 pnpchnggtc 733
 |||||

RESULT 19
 W40827
 ID W40827 standard; Protein: 1208 AA.
 XX
 AC W40827;
 XX
 XX 21-MAY-1998 (first entry)
 DT
 XX Human Jagged protein.
 DE
 XX Jagged; Notch; angiogenesis; endothelial cell; migration; human;
 KW wound repair; vulnerability; injury repair; signal transduction;
 KW motor neurone disease; amyotrophic lateral sclerosis; polymyellitis;
 KW

KW diagnosis; therapy.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Peptide 1..11

FT Domain /label= Sig_peptide

FT 175..220

FT /note= "DSL (Delta, Serrate, Lag-2 and Apx-1)

FT domain"

FT Region 224..852

FT /note= "EGF-like repeat region containing 16

FT EGF repeats"

FT Misc-difference 526

FT /note= "encoded by ANC"

FT Region 853..992

FT /note= "cysteine-rich region"

FT Domain 1058..1083

FT /note= "transmembrane domain"

FT 1084..1208

FT /note= "cytoplasmic region"

FT W09745143-A1.

PN 04-DEC-1997.

XX 30-MAY-1997; 97WO-US09407.

XX 31-MAY-1996; 96US-0018841.

XX (NAM-) NAT AMERICAN RED CROSS.

PA (UYGE-) UNIV GENEVE.

XX Maciag T, Montesano R, Pepper M, Wong MK, Zimrin AB;

XX WPI; 1998-032340/03.

XX N-PSDB; V03674.

XX New human Jagged protein - used to inhibit or promote angiogenesis
PT and to control migration of endothelial cells in injured blood
PT vessels

XX Claim 2; Page 54-61; 81pp; English.

XX This sequence comprises the human homologue of the rat Jagged
CC protein. Jagged is able to bind Notch protein and is involved in
CC endothelial cell (EC) migration and differentiation. The human
CC Jagged amino acid sequence was deduced from a human endothelial
CC cell cDNA (see V03674) induced by exposure to fibrin. Jagged
CC polypeptides can be expressed in host cell systems. A method for
CC treating or preventing disease by administering an agent that
CC (ant)agonises, inhibits, prevents, enhances or stimulates function
CC of the Notch or Jagged proteins is claimed, as well as a method for
CC affecting differentiation of mesoderm, endoderm, ectoderm and/or
CC neuroderm cells. When Jagged is applied to a micro-diameter blood
CC vessel from which ECs have been removed, damaged or reduced, it
CC decrease migrations of EC to the site, but when delivered to a
CC similar site on a large vessel it increases EC migration. Jagged
CC and its agonists are used to inhibit or prevent angiogenesis (where
CC associated with solid tumours, rheumatoid arthritis, inflammation,
CC or restenosis, particularly preventing angiogenesis from the vaso
CC vasorum and promoting large vessel EC migration to repair the lumen
CC of large vessels). Anti-Jagged and Jagged antagonists (e.g.
CC antisense Jagged and Jagged mutants) are used to promote or enhance
CC angiogenesis, particularly for wound and injury repair, e.g. where
CC surgical, traumatic and/or caused by disease, e.g. diabetes-related
CC (all claimed). Angiogenesis can be modulated in vitro or in vivo
CC and expression of proteins by gene therapy is included. Modulation
CC of the Notch-Jagged signalling pathway may also be involved in
CC placental development and motor neurone diseases such as
CC amyotrophic lateral sclerosis, poliomyelitis etc.

XX Sequence 1208 AA;

Query Match 4.5%; Score 10; DB 19; Length 1208;
Best Local Similarity 100.0%; Pred. No. 0.42;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 PNPCHNGGTC 89

|||||

Db 740 pnpchnggtc 749

RESULT 20

W05833

ID W05833 standard; Protein: 1218 AA.

XX AC W05833;

XX DT 28-JAN-1997 (first entry)

XX DE Human Serrate-1 (HJ1).

XX KW Serrate-1; human jagged-1; HJ1; Notch; cell differentiation;

KW cell fate; central nervous system; cancer; tissue repair; therapy;

KW diagnosis; antibody.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Domain 1..1067

FT Peptide /label= Extracellular_domain

FT 14..29

FT Domain /label= Sig_peptide

FT 185..229

FT /note= DSL

FT /note= "region of homology with Drosophila Delta

FT and Serrate, predicted to mediate binding

FT with Notch"

FT 234..896

FT /label= ELR

FT /note= "epidermal growth factor-like repeat domain"

FT 234..264

FT Region /label= ELR1

FT 265..299

FT Region /label= ELR2

FT 300..339

FT Region /label= ELR3

FT 340..377

FT Region /label= ELR4

FT 378..415

FT Region /label= ELR5

FT 416..453

FT Region /label= ELR6

FT 454..490

FT Region /label= ELR7

FT 491..528

FT Region /label= ELR8

FT 529..566

FT Region /label= ELR9

FT 567..598

FT /label= Partial_ELR

FT 599..632

FT /label= Partial_ELR

FT 633..670

FT /label= ELR10

FT 671..708

FT /label= ELR11

FT 709..747

FT /label= ELR12

FT 748..785

FT /label= ELR13

FT 786..823

FT /label= ELR14

FT 824..862

```
FT Region /label= ELR15
FT 863..879
FT /label= Partial_ELR
FT 880..896
FT /label= Partial_ELR
FT 1068..1089
FT /label= Transmembrane_domain
FT 1090..1218
FT /label= Intracellular_domain
XX
XX WO9627610-A1.
XX
XX 12-SEP-1996.
XX
XX 07-MAR-1996; 96WO-US03172.
XX
XX 07-MAR-1995; 95US-0400159.
XX
XX (IMCR ) IMPERIAL CANCER RES TECHNOLOGY.
XX (UYA ) UNIV YALE.
XX Artavanis-Tsakonas S, Gray GE, Henrique DMP, Ish-Horowicz D;
XX Lewis JH, Mann RS, Myat AM;
XX WPI; 1996-425379/42.
XX N-PSDB; T40090.
XX
XX Vertebrate Serrate protein and related DNA - used to treat or
XX prevent malignancies characterised by increased Notch activity.
XX
XX Claim 4; Page 95-98; 161pp; English.
XX
XX Human Serrate-1 (W05833) and human Serrate-2 (W05833) are ligands
XX for the zygotic neurogenic locus Notch, and are believed to play a
XX major role in determining cell fates (differentiation) in the
XX central nervous system. Their amino acid sequences were deduced
XX from cDNA clones (see also W40090-91) isolated from human foetal
XX brain cDNA libraries. The proteins, antibodies raised to them,
XX and encoding nucleic acids can be used in the detection of
XX Serrate sequences and in the treatment of disorders of cell fate
XX or differentiation, partic. cancer, nervous system disorders
XX and in tissue repair or regeneration.
XX
XX Sequence 1218 AA;

Query Match 4.5%; Score 10; DB 17; Length 1218;
Best Local Similarity 100.0%; Pred. No. 0.43; 0; Gaps 0;
Matches 10; Conservative 0; Mismatches 0; Indels 0;

Oy 80 PNPCHNGGTC 89
Db 750 pnpchnggtc 759

RESULT 21
W18354
ID W18354 standard; protein; 1218 AA.
XX
XX W18354;
XX
XX 11-FEB-1998 (first entry)
XX
XX Proliferation and differentiation suppression polypeptide.
XX
XX Proliferation; differentiation; suppression; human; delta-1;
XX serrate-1; blood cell; neuron; leukaemia; malignant tumour;
XX immunosuppression.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Peptide 1..31
FT
```

```
FT Protein /label= Signal
FT 32..1218
FT /label= Differentiation_suppression_protein
XX
XX WO9719172-A1.
XX
XX 29-MAY-1997.
XX
XX 15-NOV-1996; 96WO-JP03356.
XX
XX 30-NOV-1995; 95JP-0311811.
XX 17-NOV-1995; 95JP-0299611.
XX
XX (ASAH ) ASahi KASEI KOGYO KK.
XX
XX Itoh A, Sakano S;
XX WPI; 1997-298110/27.
XX N-PSDB; T70175.
XX
XX Peptide(s) encoded by human genes: delta-1 and serrate-1 - suppress
XX proliferation and differentiation of undifferentiated human blood
XX cells
XX
XX Claim 15; Page 83-91; 114pp; Japanese.
XX
XX The present sequence represents a polypeptide which suppresses
XX proliferation and differentiation of undifferentiated cells such
XX as neurons and blood cells. The polypeptide may be used for the
XX prevention and control of disorders involving undifferentiated
XX cells, such as leukaemia and malignant tumours, and improvement of
XX blood formation, e.g. after immunosuppression.
XX
XX Sequence 1218 AA;

Query Match 4.5%; Score 10; DB 18; Length 1218;
Best Local Similarity 100.0%; Pred. No. 0.43; 0; Gaps 0;
Matches 10; Conservative 0; Mismatches 0; Indels 0;

Oy 80 PNPCHNGGTC 89
Db 750 pnpchnggtc 759

RESULT 22
W44301
ID W44301 standard; Protein; 1218 AA.
XX
XX W44301;
XX
XX 19-JUN-1998 (first entry)
XX
XX Human serrate 1.
XX
XX Human; serrate 2; regulation; stem cell; differentiation; neoplasm;
XX leukaemia; endothelial cell; tumour.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Peptide 1..31
XX /label= Signal
XX Protein 32..1218
XX /label= Serrate-1
XX
XX WO9802458-A1.
XX
XX 22-JAN-1998.
XX
XX 11-JUL-1997; 97WO-JP02414.
XX
XX 14-MAY-1997; 97JP-0124063.
XX
XX
```

PR 16-JUL-1996; 96JP-0186220.
 XX (ASAH) ASahi KASEI KOGYO KK.
 XX Itoh A, Sakano S;
 XX WPI; 1998-110528/10.
 DR N-PSDB; V15201.
 XX
 XX Human serrate-2 gene expression products - used to regulate stem
 PT cell differentiation, useful in treating neoplasms, e.g. leukaemia
 PT
 XX Disclosure; Page 77-86; 103pp; Japanese.
 XX
 CC The present sequence represents human serrate 1, from the present
 CC invention which describes human serrate 2. The present invention also
 CC describes a method for the preparation of the polypeptides, and
 CC antibodies binding to the polypeptide and its fragments. The polypeptide
 CC and its fragments expressed by the serrate-2-gene can be used to inhibit
 CC stem (especially blood stem) cell differentiation and to inhibit
 CC endothelial cell growth. They may be incorporated in a cell culture
 CC media for culturing undifferentiated stem cells. They can also be used
 CC for treatment of neoplasms such as leukaemia. The antibodies can be used
 CC for the diagnosis of malignant tumours.

XX Sequence 1218 AA;

Query Match 4.5%; Score 10; DB 19; Length 1218;
 Best Local Similarity 100.0%; Pred. No. 0.43;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 PNPCHNGGTC 89
 DB 750 pnpchnggtc 759
 |||||

RESULT 23

ID W87894 standard; Protein; 1218 AA.
 XX

AC W87894;

DT 26-APR-1999 (first entry)

DE Human JAGGED1 protein.

XX JAGGED; JAGGED1; human; notch ligand; stem cell;
 KW progenitor cell; haematopoiesis; cell differentiation;
 KW Alagille syndrome; leukaemia; lymphoma; diagnosis; therapy.

OS Homo sapiens.

XX Key Location/Qualifiers
 FH Peptide 1..21
 FT /note= "signal peptide"
 FT Protein 22..1218
 FT /note= "mature protein"
 FT Domain 185..239
 FT /note= "Delta/Serrate/Lag-2 (DSL) domain"
 FT Region 234..862
 FT /note= "EGF-like repeat region"
 FT Region 863..1012
 FT /note= "cysteine-rich region"
 FT Domain 1077..1091
 FT /note= "transmembrane domain"
 FT Peptide 188..204
 FT /note= "this peptide is specifically claimed in
 Claim 1"
 FT Peptide 178..240
 FT /note= "this soluble peptide is specifically
 claimed in Claim 8"
 FT Protein 1..1010

FT /note= "this soluble protein is specifically
 FT claimed in Claim 8"
 PN WO9858958-A2.
 XX 30-DEC-1998.
 XX
 XX 25-JUN-1998; 98WO-US13207.
 XX
 XX 25-JUN-1997; 97US-0882046.
 XX (CHIL-) CHILDRENS HOSPITAL PHILADELPHIA.
 PA (UNIW) UNIV WASHINGTON.
 XX
 XX Hood L, Krantz ID, Li L, Spinner NB;
 XX WPI; 1999-081220/07.
 DR N-PSDB; V63753.

XX New Jagged peptides for inhibiting differentiation of progenitor
 PT cells - also used for maintaining these cells in undifferentiated
 PT state, e.g. for haematopoietic reconstitution
 XX

PS Claim 6; Fig 1A; 101pp; English.

XX This is the amino acid sequence of human JAGGED1 (hJAGGED1), an
 CC activating ligand for Notch protein. hJAGGED1 is expressed in
 CC bone marrow stromal cells, and a stromal cell line expressing
 CC hJAGGED1 permits survival and proliferation of haematopoietic
 CC progenitor cells expressing Notch but inhibits granulocyte
 CC differentiation. A cDNA clone (see V63753) encoding hJAGGED1
 CC was obtained from a human bone marrow cDNA library. hJAGGED1
 CC polypeptides and biologically active peptides (see W87896-98) are
 CC able (i) to inhibit differentiation of haematopoietic progenitor
 CC cells (HPC), e.g. for subsequent production of blood cells for
 CC transplantation or dendritic cells for immunotherapy, and (ii) to
 CC maintain HPC in the undifferentiated state, particularly totipotent
 CC cells or cells able to reconstitute the haematopoietic system, e.g.
 CC in patients with leukaemia or lymphoma. Treated HPC, e.g. where
 CC taken from a neonate, may be cryopreserved for many years, then
 CC thawed for further expansion and differentiation. Optionally
 CC Jagged may be provided by cells transformed to express the
 CC membrane-bound protein. Antibodies raised against hJAGGED1 can
 CC be used in a method of diagnosing Alagille syndrome by detecting
 CC reduced expression of hJAGGED1 or expression of abnormal hJAGGED1.

XX Sequence 1218 AA;

Query Match 4.5%; Score 10; DB 20; Length 1218;

Best Local Similarity 100.0%; Pred. No. 0.43;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 PNPCHNGGTC 89
 DB 750 pnpchnggtc 759
 |||||

RESULT 24

ID Y59597 standard; Protein; 1218 AA.

XX Y59597;

DT 05-APR-2000 (first entry)

DE Human Serrate protein sequence.

XX Serrate; cancerous condition; melanoma; lung cancer; breast cancer;
 KW nervous system disorder; infection; nutritional disease; therapy;
 KW cell proliferation promoter; tissue regeneration; human.
 XX Homo sapiens.

```

XX  US6004924-A.
PN  XX
PD  XX
PD  21-DEC-1999.
XX  XX
XX  06-MAR-1996; 96US-0611729.
XX  XX
XX  11-DEC-1991; 91US-0808458.
PR  14-SEP-1993; 93US-0121979.
PR  07-JUN-1994; 94US-0255102.
PR  07-MAR-1995; 95US-0400159.
XX  XX
PA  (IMCR ) IMPERIAL CANCER RES TECHNOLOGY.
PA  (UYA ) UNIV YALE.
XX  XX
PI  Ish-Horowitz D, Henrique DMP, Myat AM, Fleming RJ;
PI  Artavanis-Tsakonas S, Gray GE, Mann RS, Lewis JH;
XX  XX
DR  N-PSDB; 249096.
XX  XX
XX  Purified serrate proteins useful for treating neoplasias, nervous
PT  disorders and for promoting cell proliferation and tissue regeneration
PT  and repair.
XX  XX
XX  Claim 1; Fig 9; 114pp; English.
XX  XX
CC  This sequence represents a human serrate protein.
CC  The invention relates to purified vertebrate (mouse, chick, and human)
CC  serrate proteins. The serrate proteins may be administered to treat a
CC  cancerous condition (e.g. melanoma, lung cancer and breast cancer) by
CC  preventing progression from a preneoplastic or nonmalignant state into a
CC  neoplastic or malignant state. It may also be used to treat nervous
CC  system disorders (such as lesions caused by infections, nutritional
CC  disease and toxic substances) and to promote cell proliferation and
CC  tissue regeneration and repair. The protein itself is administered to
CC  supplement a patient's own production of serrate proteins (if levels of
CC  expression are low) or to compensate for expression of inactive proteins
CC  due to genetic mutations. The protein may also be used in the production
CC  of antibodies against serrate proteins which may be used to either down
CC  regulate serrate activity or to detect serrate proteins in samples (for
CC  example via enzyme-linked immunosorbant assay (ELISA)). The proteins may
CC  also be used to study serrate expression and its role in metabolism and
CC  to assay for agents which modulate its expression and activity.
XX  XX
SQ  Sequence 1218 AA;

Query Match 4.5%; Score 10; DB 21; Length 1218;
Best Local Similarity 100.0%; Pred. No. 0.43;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 PNPCHNGGTC 89
Db 750 pnpchnggtc 759
|||||

RESULT 25
R38304
ID R38304 standard; Protein; 1404 AA.
XX  XX
AC R38304;
XX  XX
XX  30-NOV-1993 (first entry)
XX  XX
XX  Sequence of a serrate protein.
XX  XX
KW Serrate; toporythmic protein; family.
XX  XX
OS Drosophila melanogaster.
XX  XX
PN WO9312141-A.
XX  XX

```

```

PD 24-JUN-1993.
XX  XX
XX  11-DEC-1991; 91WO-US09240.
XX  XX
PR  11-DEC-1991; 91WO-US09240.
XX  XX
PA (UYA ) UNIV YALE.
XX  XX
PI Artavanis-Tsakonas S, Fleming RJ;
XX  XX
DR WPI; 1993-214095/26.
DR N-PSDB; Q43910.
XX  XX
PT Purified serrate protein, nucleic acid and antibodies - used in
PT the study and manipulation of differentiation and other
PT physiological processes
XX  XX
XX Claim 4; Pages 74-80; 119pp; English.
XX  XX
CC Two Drosophila genomic phage libraries were screened and recombinant
CC clones were isolated. The cDNAs in lambda gt10 were isolated from an
CC early pupal library. The C1 cDNA was isolated from an early pupal
CC library. Subsequently the C3 cDNA was isolated using the 5', 700 bp
CC terminal fragment of the C1 cDNA as probe. The complete 5561bp
CC sequence of DNA of the Drosophila serrate protein was derived from
CC C1 and C3 cDNAs (Q43910). The deduced protein product appears to be
CC a transmembrane protein. AAs 51-80 represent the likely signal
CC peptide; aas 542-564 represent potential membrane associated region;
CC aas 1221-1245 represent the putative transmembrane domain.
XX  XX
SQ Sequence 1404 AA;

Query Match 4.1%; Score 9; DB 14; Length 1404;
Best Local Similarity 100.0%; Pred. No. 4.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 82 PCHNGGTC 90
Db 883 pchnggtc 891
|||||

Search completed: May 23, 2001, 06:21:20
Job time: 439 sec

```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 06:23:34 ; Search time 76.77 Seconds
(without alignments)
463.057 Million cell updates/sec

Title: US-09-237-981-29

Perfect score: 221

Sequence: 1 MKHLVAALLVGLSLGVPQF.....INAWTAAENDRPWQIVTVG 221

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 1009251 seqs, 160854530 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1009251

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 50 summaries

Database : Pending_Patents_AA_Main:*

1: /cgn2_6/ptodata/2/paa/PCTUS_COMB.pep.*
2: /cgn2_6/ptodata/2/paa/US06_COMB.pep.*
3: /cgn2_6/ptodata/2/paa/US07_COMB.pep.*
4: /cgn2_6/ptodata/2/paa/US080_COMB.pep.*
5: /cgn2_6/ptodata/2/paa/US081_COMB.pep.*
6: /cgn2_6/ptodata/2/paa/US082_COMB.pep.*
7: /cgn2_6/ptodata/2/paa/US083_COMB.pep.*
8: /cgn2_6/ptodata/2/paa/US084_COMB.pep.*
9: /cgn2_6/ptodata/2/paa/US085_COMB.pep.*
10: /cgn2_6/ptodata/2/paa/US086_COMB.pep.*
11: /cgn2_6/ptodata/2/paa/US087_COMB.pep.*
12: /cgn2_6/ptodata/2/paa/US088_COMB.pep.*
13: /cgn2_6/ptodata/2/paa/US089_COMB.pep.*
14: /cgn2_6/ptodata/2/paa/US090_COMB.pep.*
15: /cgn2_6/ptodata/2/paa/US091_COMB.pep.*
16: /cgn2_6/ptodata/2/paa/US092_COMB.pep.*
17: /cgn2_6/ptodata/2/paa/US093_COMB.pep.*
18: /cgn2_6/ptodata/2/paa/US094_COMB.pep.*
19: /cgn2_6/ptodata/2/paa/US095_COMB.pep.*
20: /cgn2_6/ptodata/2/paa/US096_COMB.pep.*
21: /cgn2_6/ptodata/2/paa/US097_COMB.pep.*
22: /cgn2_6/ptodata/2/paa/US098_COMB.pep.*
23: /cgn2_6/ptodata/2/paa/US060_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	221	100.0	221	16	US-09-237-981-29
2	194	87.8	480	8	US-08-480-229B-10
3	194	87.8	480	10	US-08-659-235B-10
4	194	87.8	480	16	US-09-237-981-10
5	45	20.4	185	23	US-60-230-435-1806
6	45	20.4	203	23	US-60-233-644-77
7	45	20.4	481	10	US-08-659-235B-29
8	45	20.4	513	8	US-08-480-229B-14
9	45	20.4	513	10	US-08-659-235B-14
10	45	20.4	513	16	US-09-237-981-14

11	41	18.6	43	8	US-08-480-229B-24	Sequence 24, Appl
12	41	18.6	43	10	US-08-659-235B-24	Sequence 24, Appl
13	41	18.6	43	16	US-09-237-981-24	Sequence 24, Appl
14	39	17.6	449	23	US-60-207-315-404	Sequence 404, App
15	31	14.0	311	23	US-60-230-435-1519	Sequence 1519, App
16	31	14.0	362	23	US-60-207-315-523	Sequence 523, App
17	27	12.2	29	23	US-60-160-203-4296	Sequence 4296, App
18	27	12.2	37	23	US-60-236-359-18942	Sequence 18942, App
19	23	10.4	62	23	US-60-192-739-3075	Sequence 3075, App
20	22	10.0	48	23	US-60-236-359-20190	Sequence 20190, App
21	21	9.5	42	23	US-60-169-840-5748	Sequence 5748, App
22	19	8.6	32	23	US-60-160-203-3897	Sequence 3897, App
23	18	8.1	25	10	US-08-480-229B-22	Sequence 22, Appl
24	18	8.1	25	16	US-08-659-235B-22	Sequence 22, Appl
25	18	8.1	25	16	US-09-237-981-22	Sequence 22, Appl
26	18	8.1	42	8	US-08-480-229B-25	Sequence 25, Appl
27	18	8.1	42	10	US-08-659-235B-25	Sequence 25, Appl
28	18	8.1	42	16	US-09-237-981-25	Sequence 25, Appl
29	18	8.1	321	8	US-08-480-229B-21	Sequence 21, Appl
30	18	8.1	321	10	US-08-659-235B-21	Sequence 21, Appl
31	18	8.1	321	16	US-09-237-981-21	Sequence 21, Appl
32	17	7.7	34	23	US-60-177-646-2311	Sequence 2311, App
33	16	7.2	44	23	US-60-177-646-2310	Sequence 2310, App
34	14	6.3	57	8	US-08-480-229B-23	Sequence 23, Appl
35	14	6.3	57	10	US-08-659-235B-23	Sequence 23, Appl
36	14	6.3	57	16	US-09-237-981-23	Sequence 23, Appl
37	10	4.5	32	23	US-60-160-203-3568	Sequence 3568, App
38	10	4.5	66	23	US-60-170-373-3161	Sequence 4202, App
39	10	4.5	77	23	US-60-170-373-3161	Sequence 4202, App
40	10	4.5	192	23	US-60-207-214-575	Sequence 575, App
41	10	4.5	249	23	US-60-213-846-540	Sequence 540, App
42	10	4.5	249	23	US-60-205-420-251	Sequence 251, App
43	10	4.5	460	23	US-60-205-420-250	Sequence 250, App
44	10	4.5	494	23	US-60-200-391-82	Sequence 82, Appl
45	10	4.5	506	23	US-60-208-129-160	Sequence 160, App
46	10	4.5	506	23	US-60-208-973-195	Sequence 195, App
47	10	4.5	673	23	US-60-201-718-120	Sequence 120, App
48	10	4.5	710	23	US-60-212-655-697	Sequence 697, App
49	10	4.5	807	23	US-60-201-718-119	Sequence 119, App
50	10	4.5	851	23	US-60-205-420-249	Sequence 249, App

ALIGNMENTS

RESULT 1
US-09-237-981-29
; Sequence 29, Application US/09237981
; GENERAL INFORMATION:
; APPLICANT: Quettermous, Thomas
; APPLICANT: Hogan, Brigid
; APPLICANT: Snodgrass, H. Ralph
; APPLICANT: Zupancic, Thomas J.
; TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
; TITLE OF INVENTION: CELL LOCUS-1
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/237,981
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:

```
; APPLICATION NUMBER: US/08/659,235
; FILING DATE: 05-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-0034-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 Pennie
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 221 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: Internal
; US-09-237-981-29

Query Match 100.0%; Score 221; DB 16; Length 221;
Best Local Similarity 100.0%; Pred. No. 2.1e-224;
Matches 221; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKHLVAALLVGLSLGVPQFGKGDICNPNCENGICLSGLADDSFSCPCPGFAGPNC 60
DB 1 MKHLVAALLVGLSLGVPQFGKGDICNPNCENGICLSGLADDSFSCPCPGFAGPNC 60
QY 61 SVVEVASDEEKPSTAGPCIPNPNCHNGTCEISEAYRGDTFYGVCPCPGFNGIHCOHNI 120
DB 61 SVVEVASDEEKPSTAGPCIPNPNCHNGTCEISEAYRGDTFYGVCPCPGFNGIHCOHNI 120
QY 121 NECEAEPCRNNGICTDLVANYSCPCPGFMRNCOYKCSGHLGIEGGIISNQIITASSNH 180
DB 121 NECEAEPCRNNGICTDLVANYSCPCPGFMRNCOYKCSGHLGIEGGIISNQIITASSNH 180
QY 181 RALFGLQKWYPYAAALNKKGLINAWTAANDRPWTQVTVG 221
DB 181 RALFGLQKWYPYAAALNKKGLINAWTAANDRPWTQVTVG 221

RESULT 2
US-08-480-229B-10
; Sequence 10, Application US/08480229B
; GENERAL INFORMATION:
; APPLICANT: Quertermous, Thomas
; APPLICANT: Hogan, Brigid
; APPLICANT: Snodgrass, H. Ralph
; APPLICANT: Zupancic, Thomas J.
; TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,229B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-0034
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 Pennie
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 221 amino acids
; TYPE: amino acid
; REFERENCE/DOCKET NUMBER: 8907-026

; APPLICATION NUMBER: US/08/659,235
; FILING DATE: 05-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-0034-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 Pennie
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 480 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-480-229B-10

Query Match 87.8%; Score 194; DB 8; Length 480;
Best Local Similarity 100.0%; Pred. No. 1.2e-195;
Matches 194; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKHLVAALLVGLSLGVPQFGKGDICNPNCENGICLSGLADDSFSCPCPGFAGPNC 60
DB 1 MKHLVAALLVGLSLGVPQFGKGDICNPNCENGICLSGLADDSFSCPCPGFAGPNC 60
QY 61 SVVEVASDEEKPSTAGPCIPNPNCHNGTCEISEAYRGDTFYGVCPCPGFNGIHCOHNI 120
DB 61 SVVEVASDEEKPSTAGPCIPNPNCHNGTCEISEAYRGDTFYGVCPCPGFNGIHCOHNI 120
QY 121 NECEAEPCRNNGICTDLVANYSCPCPGFMRNCOYKCSGHLGIEGGIISNQIITASSNH 180
DB 121 NECEAEPCRNNGICTDLVANYSCPCPGFMRNCOYKCSGHLGIEGGIISNQIITASSNH 180
QY 181 RALFGLQKWYPYAA 194
DB 181 RALFGLQKWYPYAA 194

RESULT 3
US-08-659-235B-10
; Sequence 10, Application US/08659235B
; GENERAL INFORMATION:
; APPLICANT: Quertermous, Thomas
; APPLICANT: Hogan, Brigid
; APPLICANT: Snodgrass, H. Ralph
; APPLICANT: Zupancic, Thomas J.
; TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/659,235B
; FILING DATE: 05-JUN-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-034
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 Pennie
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 480 amino acids
; TYPE: amino acid
```



```
Query Match          20.4%; Score 45; DB 23; Length 203;
Best Local Similarity 100.0%; Pred. No. 6.1e-39;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 PNPCHNGGTCEISEAYRGDTFTGYGCKCPRGFNGIHCOHNECE 124
|||||
DB 58 PNPCHNGGTCEISEAYRGDTFTGYGCKCPRGFNGIHCOHNECE 102
|||||

RESULT 7
US-08-659-235B-29
; Sequence 29, Application US/08659235B
; GENERAL INFORMATION:
; APPLICANT: Quertermous, Thomas
; APPLICANT: Hogan, Brigid
; APPLICANT: Snodgrass, H. Ralph
; APPLICANT: Zupancic, Thomas J.
; TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2711
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/659,235B
; FILING DATE: 05-JUN-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-034
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 Pennie
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 481 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-659-235B-29

Query Match          20.4%; Score 45; DB 10; Length 481;
Best Local Similarity 100.0%; Pred. No. 1.4e-38;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 PNPCHNGGTCEISEAYRGDTFTGYGCKCPRGFNGIHCOHNECE 124
|||||
DB 80 PNPCHNGGTCEISEAYRGDTFTGYGCKCPRGFNGIHCOHNECE 124
|||||

RESULT 8
US-08-480-229B-14
; Sequence 14, Application US/08480229B
; GENERAL INFORMATION:
; APPLICANT: Quertermous, Thomas
; APPLICANT: Hogan, Brigid
; APPLICANT: Snodgrass, H. Ralph
; APPLICANT: Zupancic, Thomas J.
; TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
```

```
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2711
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,229B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-026
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 Pennie
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 513 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-480-229B-14

Query Match          20.4%; Score 45; DB 8; Length 513;
Best Local Similarity 100.0%; Pred. No. 1.5e-38;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 PNPCHNGGTCEISEAYRGDTFTGYGCKCPRGFNGIHCOHNECE 124
|||||
DB 112 PNPCHNGGTCEISEAYRGDTFTGYGCKCPRGFNGIHCOHNECE 156
|||||

RESULT 9
US-08-659-235B-14
; Sequence 14, Application US/08659235B
; GENERAL INFORMATION:
; APPLICANT: Quertermous, Thomas
; APPLICANT: Hogan, Brigid
; APPLICANT: Snodgrass, H. Ralph
; APPLICANT: Zupancic, Thomas J.
; TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2711
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/659,235B
; FILING DATE: 05-JUN-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
```


REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-034
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennie
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 513 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-659-235B-14

Query Match 20.4%; Score 45; DB 10; Length 513;
Best Local Similarity 100.0%; Pred. No. 1.5e-38;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 80 PNPCHNGTCEISEAYRGDTFIGYVCKCPRGFNGIHCOHNECE 124
|||||
Db 112 PNPCHNGTCEISEAYRGDTFIGYVCKCPRGFNGIHCOHNECE 156

RESULT 10
US-09-237-981-14
Sequence 14, Application US/09237981
GENERAL INFORMATION:
APPLICANT: Quertermous, Thomas
APPLICANT: Hogan, Brigid
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/237,981
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/659,235
FILING DATE: 05-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-0034-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennie
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 513 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-237-981-14

Query Match 20.4%; Score 45; DB 16; Length 513;
Best Local Similarity 100.0%; Pred. No. 1.5e-38;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 PNPCHNGTCEISEAYRGDTFIGYVCKCPRGFNGIHCOHNECE 124
|||||
Db 112 PNPCHNGTCEISEAYRGDTFIGYVCKCPRGFNGIHCOHNECE 156

RESULT 11
US-08-480-229B-24
Sequence 24, Application US/08480229B
GENERAL INFORMATION:
APPLICANT: Quertermous, Thomas
APPLICANT: Hogan, Brigid
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,229B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-026
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennie
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 43 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-480-229B-24
Query Match 18.6%; Score 41; DB 8; Length 43;
Best Local Similarity 100.0%; Pred. No. 2.4e-35;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 80 PNPCHNGTCEISEAYRGDTFIGYVCKCPRGFNGIHCOHNI 120
|||||
Db 3 PNPCHNGTCEISEAYRGDTFIGYVCKCPRGFNGIHCOHNI 43
RESULT 12
US-08-659-235B-24
Sequence 24, Application US/08659235B
GENERAL INFORMATION:
APPLICANT: Quertermous, Thomas
APPLICANT: Hogan, Brigid
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
CELL LOCUS-1

NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/659,235B
FILING DATE: 05-JUN-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-034
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 43 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-659-235B-24

Query Match 18.6%; Score 41; DB 10; Length 43;
Best Local Similarity 100.0%; Pred. No. 2.4e-35;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 PNPCHNGGTCEISEAYRGDTFGYVCKCPRGFNGIHCOHNI 120
|||||
Db 3 PNPCHNGGTCEISEAYRGDTFGYVCKCPRGFNGIHCOHNI 43

RESULT 13
US-09-237-981-24
Sequence 24, Application US/09237981
GENERAL INFORMATION:
APPLICANT: Quentermous, Thomas
APPLICANT: Hogan, Brigid
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
TITLE OF INVENTION: CELL LOCUS-1
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/237,981
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/659,235

FILING DATE: 05-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-0034-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennie
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 43 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-09-237-981-24

Query Match 18.6%; Score 41; DB 16; Length 43;
Best Local Similarity 100.0%; Pred. No. 2.4e-35;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 PNPCHNGGTCEISEAYRGDTFGYVCKCPRGFNGIHCOHNI 120
|||||
Db 3 PNPCHNGGTCEISEAYRGDTFGYVCKCPRGFNGIHCOHNI 43

RESULT 14
US-60-207-315-404
Sequence 404, Application US/60207315
GENERAL INFORMATION:
APPLICANT: Beasley, Ellen
TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
TITLE OF INVENTION: USES THEREOF
FILE REFERENCE: CL000601
CURRENT APPLICATION NUMBER: US/60/207,315
CURRENT FILING DATE: 2000-05-30
NUMBER OF SEQ ID NOS: 528
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 404
LENGTH: 449
TYPE: PRT
ORGANISM: HUMAN
FEATURE:
NAME/KEY: VARIANT
LOCATION: (1)...(449)
OTHER INFORMATION: Xaa = Any Amino Acid
US-60-207-315-404

Query Match 17.6%; Score 39; DB 23; Length 449;
Best Local Similarity 100.0%; Pred. No. 2.7e-32;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 PNPCHNGGTCEISEAYRGDTFGYVCKCPRGFNGIHCOH 118
|||||
Db 5 PNPCHNGGTCEISEAYRGDTFGYVCKCPRGFNGIHCOH 43

RESULT 15
US-60-230-435-1519
Sequence 1519, Application US/60230435
GENERAL INFORMATION:
APPLICANT: Beasley, Ellen
TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
TITLE OF INVENTION: USES THEREOF
FILE REFERENCE: CL000768
CURRENT APPLICATION NUMBER: US/60/230,435
CURRENT FILING DATE: 2000-09-06
NUMBER OF SEQ ID NOS: 2991


```

Query Match          10.4%; Score 23; DB 23; Length 62;
Best Local Similarity 100.0%; Pred. No. 3.le-16;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 196 LNKKGLINAWTAAENDRPWQIV 218
DB 40 LNKKGLINAWTAAENDRPWQIV 62

RESULT 20
US-60-236-359-20190
; Sequence 20190, Application US/60236359
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: MDHMORF-4P
; CURRENT APPLICATION NUMBER: US/60/236.359
; CURRENT FILING DATE: 2000-05-27
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; NUMBER OF SEQ ID NOS: 21709
; SOFTWARE: Molecular Dynamics Sequence Listing Engine
; SEQ ID NO 20190
; LENGTH: 48
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC008430.2
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.5
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.97
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.92
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN HEL100, SIGNAL = 0.99
; OTHER INFORMATION: EST_HUMAN HIT: AAL12613.1, EVALUE 1.00e-19
; OTHER INFORMATION: SWISSPROT HIT: P79385, EVALUE 2.00e-10
US-60-236-359-20190

```

```

Query Match          10.0%; Score 22; DB 23; Length 48;
Best Local Similarity 100.0%; Pred. No. 2.8e-15;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 196 LNKKGKGLINAWTAAENDRWPWQ 217
      |||||
Db 27 LNKKGKGLINAWTAAENDRWPWQ 48

RESULT 21
US-60-169-840-5748
; Sequence 5748, Application US/60169840
; GENERAL INFORMATION:
; APPLICANT: Bonazzi, Vivien
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL000164
; CURRENT APPLICATION NUMBER: US/60/169,840
; CURRENT FILING DATE: 1999-12-09
; NUMBER OF SEQ ID NOS: 9628
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5748
; LENGTH: 42
; TYPE: PRT
; ORGANISM: Human
US-60-169-840-5748

```

```

Query Match          9.5%; Score 21; DB 23; Length 42;
Best Local Similarity 100.0%; Pred.No. 2.8e-14;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 136 DLVANSCECPGEFMGRNCQY 156
      |||
Db 18 DLVANSCECPGEFMGRNCQY 38

RESULT 22
US-60-160-203-3897
; Sequence 3897, Application US/60160203
; GENERAL INFORMATION:
; APPLICANT: BONAZZI, VIVIEN
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CLO00116
; CURRENT APPLICATION NUMBER: US/60/160,203
; CURRENT FILING DATE: 1999-10-19
; NUMBER OF SEQ ID NOS: 6374
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3897
; LENGTH: 32
; TYPE: PRT
; ORGANISM: HUMAN
; US-60-160-203-3897

```

```

Query Match      8.6%; Score 19; DB 23; Length 32;
Best Local Similarity 100.0%; Pred. NO. 2.8e-12;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 136 DLVANYSCPCGPFMGMRNC 154
      |||||
Db 14 DLVANYSCPCGPFMGMRNC 32

RESULT 23
US-08-480-229B-22
; Sequence 22, Application US/08480229B
; GENERAL INFORMATION:
; APPLICANT: Quertermous, Thomas
; APPLICANT: Hogan, Brigid
; APPLICANT: Snodgrass, H. Ralph
; APPLICANT: Zupancic, Thomas J.
; TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
; TITLE OF INVENTION: CELL LOCUS-1
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,229B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Foissant, Brian M.
; REGISTRATION NUMBER: 28,462
; REFERENCE/POCKET NUMBER: 8907-026
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741

```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 06:23:48 ; Search time 9.39 seconds
(without alignments)
168.707 Million cell updates/sec

Title: US-09-237-981-29
Perfect score: 221
Sequence: 1 MKHLVAALLVLSLGVQPF.....INAWTAENDRWPIQVTVG 221

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 41833 seqs, 7168127 residues

Word size : 0
Total number of hits satisfying chosen parameters: 41833

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 50 summaries

Database : Pending Patents AA.New.*
1: /cgn2_6/ptodata/1/paa/PCT_NEW_COMB.pcp.*
2: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pcp.*
3: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pcp.*
4: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pcp.*
5: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pcp.*
6: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7	3.2	383	1	PCT-US01-11988-770
2	7	3.2	383	5	US-09-833-245-770
3	7	3.2	527	5	US-09-422-601-1
4	7	3.2	750	6	US-60-248-505-1003
5	7	3.2	795	6	US-60-248-505-1081
6	7	3.2	4545	5	US-09-750-972-2
7	6	2.7	14	1	PCT-US01-06436-36
8	6	2.7	14	1	PCT-US01-06436-39
9	6	2.7	66	1	PCT-US01-01238-42
10	6	2.7	81	5	US-09-739-449-10947
11	6	2.7	83	5	US-09-739-449-11088
12	6	2.7	86	5	US-09-640-211A-805
13	6	2.7	145	5	US-09-814-666-274
14	6	2.7	148	1	PCT-US01-01302-91
15	6	2.7	154	5	US-09-739-449-11785
16	6	2.7	186	1	PCT-US01-11988-924
17	6	2.7	186	5	US-09-833-245-924
18	6	2.7	188	1	PCT-US01-01302-109
19	6	2.7	205	4	US-08-706-945B-135
20	6	2.7	223	5	US-09-739-449-9657
21	6	2.7	234	5	US-09-640-211A-614
22	6	2.7	244	5	US-09-308-823A-596
23	6	2.7	245	1	PCT-US01-11988-922
24	6	2.7	245	5	US-09-833-245-922
25	6	2.7	288	1	PCT-US01-01302-107
26	6	2.7	298	5	US-09-739-449-8950
27	6	2.7	304	5	US-09-739-449-11600

28	6	2.7	327	6	US-60-282-814-5	Sequence 5, Appli
29	6	2.7	365	5	US-09-739-449-12197	Sequence 12197, A
30	6	2.7	379	5	US-09-808-689-4	Sequence 4, Appli
31	6	2.7	379	5	US-09-423-844-4	Sequence 4, Appli
32	6	2.7	379	5	US-09-403-296A-4	Sequence 4, Appli
33	6	2.7	379	5	US-09-380-139A-4	Sequence 1022, Ap
34	6	2.7	397	6	US-60-248-505-1022	Sequence 1064, Ap
35	6	2.7	400	1	PCT-US01-04098A-1064	Sequence 8652, Ap
36	6	2.7	404	5	US-09-739-449-8652	Sequence 11756, A
37	6	2.7	411	5	US-09-739-449-11756	Sequence 3032, Ap
38	6	2.7	413	1	PCT-US01-04098A-3032	Sequence 7, Appli
39	6	2.7	434	5	US-09-827-040-7	Sequence 2, Appli
40	6	2.7	476	5	US-09-827-040-2	Sequence 3, Appli
41	6	2.7	476	5	US-09-827-040-3	Sequence 6, Appli
42	6	2.7	476	5	US-09-827-040-6	Sequence 1035, Ap
43	6	2.7	491	1	PCT-US01-11988-1035	Sequence 1057, Ap
44	6	2.7	491	1	PCT-US01-11988-1057	Sequence 1035, Ap
45	6	2.7	491	5	US-09-833-245-1035	Sequence 1057, Ap
46	6	2.7	491	5	US-09-833-245-1057	Sequence 12458, A
47	6	2.7	505	5	US-09-739-449-12458	Sequence 5, Appli
48	6	2.7	561	5	US-09-827-040-5	Sequence 33, Appli
49	6	2.7	601	1	PCT-US01-08117-33	Sequence 9500, Ap
50	6	2.7	645	5	US-09-739-449-9500	

ALIGNMENTS

RESULT 1
PCT-US01-11988-770
; Sequence 770, Application PC/TUS0111988
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546PCT
; CURRENT APPLICATION NUMBER: PCT/US01/11988
; CURRENT FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: 60/229, 358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256, 931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199, 384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 770
; LENGTH: 383
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-11988-770

Query Match 3.2%; Score 7; DB 1; Length 383;
Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 CPEGFAG 56
|||||||
Db 200 CPEGFAG 206

RESULT 2
US-09-833-245-770
; Sequence 770, Application US/09833245
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546PCT
; CURRENT APPLICATION NUMBER: US/09/833,245
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229, 358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256, 931

```

; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199, 384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 770
; LENGTH: 383
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-833-245-770

Query Match      3.2%; Score 7; DB 5; Length 383;
Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 CPEGFAG 56
   |||||
Db 200 CPEGFAG 206

RESULT 3
US-09-422-601-1
; Sequence 1, Application US/09422601
; GENERAL INFORMATION:
; APPLICANT: Baizar, Laszlo S.
; APPLICANT: Nesheim, Michael R.
; APPLICANT: Church, William R.
; TITLE OF INVENTION: Compositions for Inhibiting the
; TITLE OF INVENTION: Activation of Thrombin-Activatable Fibrinolysis Inhibitor
; TITLE OF INVENTION: (TAFI)
; FILE REFERENCE: V0139/7036 (HCL/MAT)
; CURRENT APPLICATION NUMBER: US/09/422,601
; CURRENT FILING DATE: 1999-10-21
; PRIOR APPLICATION NUMBER: US 08/966,432
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: US 60/030,721
; PRIOR FILING DATE: 1996-11-08
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 527
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-422-601-1

Query Match      3.2%; Score 7; DB 5; Length 527;
Best Local Similarity 100.0%; Pred. No. 7.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 CPEGFAG 56
   |||||
Db 75 CPEGFAG 81

RESULT 4
US-60-248-505-1003
; Sequence 1003, Application US/60248505
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF
; FILE REFERENCE: c1000918
; CURRENT APPLICATION NUMBER: US/60/248,505
; CURRENT FILING DATE: 2000-11-15
; NUMBER OF SEQ ID NOS: 1998
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1003
; LENGTH: 750
; TYPE: PRT
; ORGANISM: Human

```

US-60-248-505-1003

Query Match 3.2%; Score 7; DB 6; Length 750;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 WLLVGLS 14
 |||||
 Db 411 WLLVGLS 417

RESULT 5
 US-60-248-505-1081
 ; Sequence 1081, Application US/60248505
 ; GENERAL INFORMATION:
 ; APPLICANT: Beasley, Ellen
 ; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
 ; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
 ; TITLE OF INVENTION: PROTEINS, AND USES THEREOF
 ; FILE REFERENCE: c1000918
 ; CURRENT APPLICATION NUMBER: US/60/248,505
 ; CURRENT FILING DATE: 2000-11-15
 ; NUMBER OF SEQ ID NOS: 1998
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1081
 ; LENGTH: 795
 ; TYPE: PRT
 ; ORGANISM: Human
 ; FEATURE:
 ; NAME/KEY: VARIANT
 ; LOCATION: (1)...(795)
 ; OTHER INFORMATION: Xaa = Any Amino Acid
 US-60-248-505-1081

Query Match 3.2%; Score 7; DB 6; Length 795;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 WLLVGLS 14
 |||||
 Db 416 WLLVGLS 422

RESULT 6
 US-09-750-972-2
 ; Sequence 2, Application US/09750972
 ; GENERAL INFORMATION:
 ; APPLICANT: Pramod K. Srivastava
 ; TITLE OF INVENTION: ALPHA(2) MACROGLOBULIN RECEPTOR AS A HEAT SHOCK
 ; TITLE OF INVENTION: PROTEIN RECEPTOR AND USES THEREOF
 ; FILE REFERENCE: 8449-134
 ; CURRENT APPLICATION NUMBER: US/09/750,972
 ; CURRENT FILING DATE: 2000-12-28
 ; PRIOR APPLICATION NUMBER: 09/750,972
 ; PRIOR FILING DATE: 2000-12-28
 ; PRIOR APPLICATION NUMBER: 09/668,724
 ; PRIOR FILING DATE: 2000-09-22
 ; NUMBER OF SEQ ID NOS: 57
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 2
 ; LENGTH: 4545
 ; TYPE: PRT
 ; ORGANISM: Mus musculus
 US-09-750-972-2

Query Match 3.2%; Score 7; DB 5; Length 4545;
 Best Local Similarity 100.0%; Pred. No. 56;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 83 CHNGGTC 89

Db 9 LVGLSL 14

RESULT 12

US-09-640-211A-805

; Sequence 805, Application US/09640211A
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; FILE OF INVENTION: Modification of Gene Transcription
; FILE REFERENCE: 11000.1021C1U
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 805
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Eucalyptus grandis
US-09-640-211A-805

Query Match

Best Local Similarity 2.7%; Score 6; DB 5; Length 86;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 204 AWTAAE 209

Db 16 AWTAAE 21

RESULT 13

US-09-814-666-274

; Sequence 274, Application US/09814666
; GENERAL INFORMATION:
; APPLICANT: Gearing, David P.
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A
; FILE OF INVENTION: HUMAN PROSTATE EPITHELIUM LIBRARY
; FILE REFERENCE: 1600.1018-001
; CURRENT APPLICATION NUMBER: US/09/814,666
; CURRENT FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: US 09/307,649
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: 60/084,562
; PRIOR FILING DATE: 1998-05-07
; NUMBER OF SEQ ID NOS: 394
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 274
; LENGTH: 145
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)...(22)
; NAME/KEY: VARIANT
; LOCATION: (1)...(145)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-814-666-274

Query Match

Best Local Similarity 2.7%; Score 6; DB 5; Length 145;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 154 COYKCS 159

Db 63 COYKCS 68

RESULT 14

PCT-US01-01302-91

; Sequence 91, Application PC/TUS0101302
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc., et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ07PCT
; CURRENT APPLICATION NUMBER: PCT/US01/01302
; CURRENT FILING DATE: 2001-01-17
; NUMBER OF SEQ ID NOS: 192
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 91
; LENGTH: 148
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-01302-91

Query Match

Best Local Similarity 2.7%; Score 6; DB 1; Length 148;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 PCRNGG 132

Db 50 PCRNGG 55

RESULT 15

US-09-739-449-11785

; Sequence 11785, Application US/09739449
; GENERAL INFORMATION:
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15490)C
; CURRENT APPLICATION NUMBER: US/09/739,449
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 09/514,000
; PRIOR FILING DATE: 2000-02-23
; NUMBER OF SEQ ID NOS: 13351
; SEQ ID NO 11785
; LENGTH: 154
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-09-739-449-11785

Query Match

Best Local Similarity 2.7%; Score 6; DB 5; Length 154;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 194 AALNKK 199

Db 144 AALNKK 149

RESULT 16

PCT-US01-11988-924

; Sequence 924, Application PC/TUS0111988
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546PCT
; CURRENT APPLICATION NUMBER: PCT/US01/11988
; CURRENT FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 924

; LENGTH: 186

; TYPE: PRT

; ORGANISM: Homo sapiens

PCT-US01-11988-924

Query Match

Best Local Similarity 2.7%; Score 6; DB 1; Length 186;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 13 LSLGVP 18

Db 79 LSLGVP 84

RESULT 17

US-09-833-245-924

; Sequence 924, Application US/09833245

; GENERAL INFORMATION:

; APPLICANT: Human Genome Sciences, Inc.

; TITLE OF INVENTION: Albumin Fusion Proteins

; FILE REFERENCE: PF546PCT

; CURRENT APPLICATION NUMBER: US/09/833,245

; CURRENT FILING DATE: 2001-04-12

; PRIOR APPLICATION NUMBER: 60/229, 358

; PRIOR FILING DATE: 2000-04-12

; PRIOR APPLICATION NUMBER: 60/256, 931

; PRIOR FILING DATE: 2000-12-21

; PRIOR APPLICATION NUMBER: 60/199, 384

; PRIOR FILING DATE: 2000-04-25

; NUMBER OF SEQ ID NOS: 2267

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 924

; LENGTH: 186

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-833-245-924

Query Match

Best Local Similarity 2.7%; Score 6; DB 5; Length 186;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 13 LSLGVP 18

Db 79 LSLGVP 84

RESULT 18

PCT-US01-01302-109

; Sequence 109, Application PC/TUS0101302

; GENERAL INFORMATION:

; APPLICANT: Human Genome Sciences, Inc., et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

; FILE REFERENCE: PTZ07PCT

; CURRENT APPLICATION NUMBER: PCT/US01/01302

; CURRENT FILING DATE: 2001-01-17

; NUMBER OF SEQ ID NOS: 192

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 109

; LENGTH: 188

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: SITE

; LOCATION: (97)

; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

; NAME/KEY: SITE

; LOCATION: (187)

; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

; NAME/KEY: SITE

; LOCATION: (188)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
PCT-US01-01302-109

Query Match

Best Local Similarity 2.7%; Score 6; DB 1; Length 188;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 127 PCRRGG 132

Db 82 PCRRGG 87

RESULT 19

US-08-706-945B-135

; Sequence 135, Application US/08706945B

; GENERAL INFORMATION:

; APPLICANT: Boyle, William

; APPLICANT: Lacey, David

; APPLICANT: Calzone, Frank

; APPLICANT: Chang, Ming-Shi

; TITLE OF INVENTION: Osteoprotegerin

; FILE REFERENCE: A-378CIP

; CURRENT APPLICATION NUMBER: US/08/706,945B

; CURRENT FILING DATE: 2001-04-18

; PRIOR APPLICATION NUMBER: 08/577,788

; PRIOR FILING DATE: 1995-12-22

; NUMBER OF SEQ ID NOS: 142

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 135

; LENGTH: 205

; TYPE: PRT

; ORGANISM: Homo sapiens

US-08-706-945B-135

Query Match

Best Local Similarity 2.7%; Score 6; DB 4; Length 205;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12 GLSLGV 17

Db 15 GLSLGV 20

RESULT 20

US-09-739-449-9657

; Sequence 9657, Application US/09739449

; GENERAL INFORMATION:

; APPLICANT: Hinkle, Gregory J.

; APPLICANT: Slater, Steven C.

; TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof

; FILE REFERENCE: 38-10(15490)C

; CURRENT APPLICATION NUMBER: US/09/739,449

; CURRENT FILING DATE: 2000-12-19

; PRIOR APPLICATION NUMBER: US 09/514,000

; PRIOR FILING DATE: 2000-02-23

; NUMBER OF SEQ ID NOS: 13351

; SEQ ID NO 9657

; LENGTH: 223

; TYPE: PRT

; ORGANISM: Agrobacterium tumefaciens

US-09-739-449-9657

Query Match

Best Local Similarity 2.7%; Score 6; DB 5; Length 223;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 AAWLLV 11

Db 140 AAWLLV 145

RESULT 21
US-09-640-211A-614
; Sequence 614, Application US/09640211A
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; FILE OF INVENTION: Modification of Gene Transcription
; FILE REFERENCE: 11000.1021C10
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 614
; LENGTH: 234
; TYPE: PRT
; ORGANISM: Eucalyptus grandis
US-09-640-211A-614

Query Match 2.7%; Score 6; DB 5; Length 234;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 34 GGICLS 39
Db 193 GGICLS 198

RESULT 22
US-09-308-823A-596
; Sequence 596, Application US/09308823A
; GENERAL INFORMATION:
; APPLICANT: Warren, Richard
; TITLE OF INVENTION: Novel Prokaryotic Polynucleotides,
; FILE OF INVENTION: Polypeptides and Their Uses
; FILE REFERENCE: P50596
; CURRENT APPLICATION NUMBER: US/09/308,823A
; CURRENT FILING DATE: 1999-05-25
; NUMBER OF SEQ ID NOS: 661
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 596
; LENGTH: 244
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-308-823A-596

Query Match 2.7%; Score 6; DB 5; Length 244;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 199 KGLINA 204
Db 15 KGLINA 20

RESULT 23
PCT-US01-11988-922
; Sequence 922, Application PC/TUS0111988
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; FILE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546PCT
; CURRENT APPLICATION NUMBER: PCT/US01/11988
; CURRENT FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: 60/229, 358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256, 931

; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199, 384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 922
; LENGTH: 245
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-11988-922

Query Match 2.7%; Score 6; DB 1; Length 245;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 13 LSLGVP 18
Db 104 LSLGVP 109

RESULT 24
US-09-833-245-922
; Sequence 922, Application US/09833245
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546PCT
; CURRENT APPLICATION NUMBER: US/09/833,245
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229, 358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256, 931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199, 384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 922
; LENGTH: 245
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-833-245-922

Query Match 2.7%; Score 6; DB 5; Length 245;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 13 LSLGVP 18
Db 104 LSLGVP 109

RESULT 25
PCT-US01-01302-107
; Sequence 107, Application PC/TUS0101302
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc., et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ07PCT
; CURRENT APPLICATION NUMBER: PCT/US01/01302
; CURRENT FILING DATE: 2001-01-17
; NUMBER OF SEQ ID NOS: 192
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 107
; LENGTH: 288
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-01302-107

Query Match 2.7%; Score 6; DB 1; Length 288;

Best Local Similarity 100.0%; Pred. No. 50;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 PCRNGG 132
 |||||
Db 125 PCRNGG 130

Search completed: May 23, 2001, 06:23:49
Job time: 173 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 06:21:42 ; Search time 20.57 Seconds
(without alignments)
206.398 Million cell updates/sec

Title: US-09-237-981-29
Perfect score: 221
Sequence: 1 MKHLVAALLVGLSLGVQF.....INAWTAAENDRWFIQTVG 221

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 185757 seqs, 19210857 residues

Word size : 0

Total number of hits satisfying chosen parameters: 185757

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 50 summaries

Database : Issued Patents AA: *
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	221	100.0	221	2	US-08-480-229C-29
2	221	100.0	221	2	US-08-659-235C-29
3	194	87.8	480	2	US-08-480-229C-10
4	194	87.8	480	2	US-08-659-235C-10
5	45	20.4	513	2	US-08-480-229C-14
6	45	20.4	513	2	US-08-659-235C-14
7	41	18.6	43	2	US-08-480-229C-24
8	41	18.6	43	2	US-08-659-235C-24
9	18	8.1	25	2	US-08-480-229C-22
10	18	8.1	25	2	US-08-659-235C-22
11	18	8.1	42	2	US-08-480-229C-25
12	18	8.1	42	2	US-08-659-235C-25
13	18	8.1	321	2	US-08-480-229C-21
14	18	8.1	321	2	US-08-659-235C-21
15	14	6.3	57	2	US-08-480-229C-23
16	14	6.3	57	2	US-08-659-235C-23
17	10	4.5	1010	4	US-08-882-046-7
18	10	4.5	1193	2	US-08-400-159-10
19	10	4.5	1193	3	US-08-611-729A-10
20	10	4.5	1218	2	US-08-400-159-6
21	10	4.5	1218	3	US-08-611-729A-6
22	10	4.5	1218	4	US-08-882-046-2
23	10	4.5	1219	4	US-08-882-046-5
24	10	4.5	2556	1	US-08-185-432-17
25	10	4.5	2556	1	US-08-083-590A-20
26	10	4.5	2556	3	US-08-532-384-20
27	9	4.1	1404	2	US-08-400-159-2

28	9	4.1	1404	3	US-08-611-729A-2	Sequence 2, Appli
29	8	3.6	830	4	US-08-872-855-11	Sequence 11, Appl
30	8	3.6	833	1	US-08-264-534-6	Sequence 6, Appli
31	8	3.6	833	1	US-08-083-590A-2	Sequence 2, Appli
32	8	3.6	833	1	US-08-465-500-6	Sequence 6, Appli
33	8	3.6	833	2	US-08-346-126-6	Sequence 6, Appli
34	8	3.6	833	2	US-08-346-128-6	Sequence 6, Appli
35	8	3.6	833	3	US-08-532-384-2	Sequence 2, Appli
36	8	3.6	833	3	US-08-893-828-6	Sequence 6, Appli
37	8	3.6	908	5	PCT-US95-03747-3	Sequence 3, Appli
38	8	3.6	912	5	PCT-US95-03747-2	Sequence 3, Appli
39	8	3.6	1251	5	PCT-US95-02251-3	Sequence 3, Appli
40	8	3.6	1252	1	US-08-199-780-3	Sequence 3, Appli
41	8	3.6	1252	2	US-08-316-650-3	Sequence 3, Appli
42	8	3.6	1253	3	US-08-479-722B-4	Sequence 4, Appli
43	8	3.6	2523	1	US-08-185-432-18	Sequence 18, Appli
44	7	3.2	15	1	US-08-179-481-14	Sequence 14, Appli
45	7	3.2	114	2	US-08-733-564-1	Sequence 1, Appli
46	7	3.2	115	1	US-08-312-870-9	Sequence 9, Appli
47	7	3.2	275	1	US-08-312-870-7	Sequence 7, Appli
48	7	3.2	358	3	US-09-041-545-2	Sequence 2, Appli
49	7	3.2	358	3	US-09-327-925-2	Sequence 2, Appli
50	7	3.2	387	2	US-08-162-402B-6	Sequence 6, Appli

ALIGNMENTS

RESULT 1
US-08-480-229C-29
; Sequence 29, Application US/08480229C
; Patent No. 5874562
; GENERAL INFORMATION:
; APPLICANT: Quertermous, Thomas
; APPLICANT: Hogan, Brigid
; APPLICANT: Snodgrass, H. Ralph
; APPLICANT: Zupancic, Thomas J.
; TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
; TITLE OF INVENTION: CELL LOCUS-1
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,229C
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-0026-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 Pennie
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 221 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; US-08-480-229C-29

```
Query Match      100.0%; Score 221; DB 2; Length 221;
Best Local Similarity 100.0%; Pred. No. 1.4e-207;
Matches 221; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKHLVAALLVGLSLGVDPQFGKGDICNPNCENGICLSGLADDSFSCPCPGFAGPNC 60
DB 1 MKHLVAALLVGLSLGVDPQFGKGDICNPNCENGICLSGLADDSFSCPCPGFAGPNC 60

QY 61 SVVEVASDEEKPTSGAGPCIPNPNCHNGTCEISEAYRGDTFIGVCKCPRGFNGIHCOHNI 120
DB 61 SVVEVASDEEKPTSGAGPCIPNPNCHNGTCEISEAYRGDTFIGVCKCPRGFNGIHCOHNI 120

QY 121 NECEAPECRNGICTDLVANYSCPCPGCFMGRNCOYKCSGHLGIEGGIISNOQITASSNH 180
DB 121 NECEAPECRNGICTDLVANYSCPCPGCFMGRNCOYKCSGHLGIEGGIISNOQITASSNH 180

QY 181 RALFGLQKWYPYAAALNKKGLINAWTAENDRWPWIOQTG 221
DB 181 RALFGLQKWYPYAAALNKKGLINAWTAENDRWPWIOQTG 221

RESULT 2
US-08-659-235C-29
; Sequence 29, Application US/08659235C
; Patent No. 5877281
; GENERAL INFORMATION:
; APPLICANT: Quettermous, Thomas
; APPLICANT: Hogan, Brigid
; APPLICANT: Snodgrass, H. Ralph
; APPLICANT: Zupancic, Thomas J.
; TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 05-JUN-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-0034-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 Pennie
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 221 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
US-08-659-235C-29

Query Match      100.0%; Score 221; DB 2; Length 221;
Best Local Similarity 100.0%; Pred. No. 1.4e-207;
Matches 221; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MKHLVAALLVGLSLGVDPQFGKGDICNPNCENGICLSGLADDSFSCPCPGFAGPNC 60
DB 1 MKHLVAALLVGLSLGVDPQFGKGDICNPNCENGICLSGLADDSFSCPCPGFAGPNC 60

QY 61 SVVEVASDEEKPTSGAGPCIPNPNCHNGTCEISEAYRGDTFIGVCKCPRGFNGIHCOHNI 120
DB 61 SVVEVASDEEKPTSGAGPCIPNPNCHNGTCEISEAYRGDTFIGVCKCPRGFNGIHCOHNI 120

QY 121 NECEAPECRNGICTDLVANYSCPCPGCFMGRNCOYKCSGHLGIEGGIISNOQITASSNH 180
DB 121 NECEAPECRNGICTDLVANYSCPCPGCFMGRNCOYKCSGHLGIEGGIISNOQITASSNH 180

QY 181 RALFGLQKWYPYAAALNKKGLINAWTAENDRWPWIOQTG 221
DB 181 RALFGLQKWYPYAAALNKKGLINAWTAENDRWPWIOQTG 221

RESULT 3
US-08-480-229C-10
; Sequence 10, Application US/08480229C
; Patent No. 5874562
; GENERAL INFORMATION:
; APPLICANT: Quettermous, Thomas
; APPLICANT: Hogan, Brigid
; APPLICANT: Snodgrass, H. Ralph
; APPLICANT: Zupancic, Thomas J.
; TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-0026-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 Pennie
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 480 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-480-229C-10

Query Match      87.8%; Score 194; DB 2; Length 480;
Best Local Similarity 100.0%; Pred. No. 5e-181;
Matches 194; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKHLVAALLVGLSLGVDPQFGKGDICNPNCENGICLSGLADDSFSCPCPGFAGPNC 60
DB 1 MKHLVAALLVGLSLGVDPQFGKGDICNPNCENGICLSGLADDSFSCPCPGFAGPNC 60

QY 61 SVVEVASDEEKPTSGAGPCIPNPNCHNGTCEISEAYRGDTFIGVCKCPRGFNGIHCOHNI 120
DB 61 SVVEVASDEEKPTSGAGPCIPNPNCHNGTCEISEAYRGDTFIGVCKCPRGFNGIHCOHNI 120
```


Db 61 SVVEVASDEEKPTSGPCIPNCHNGGTCEISEAYRGDTFIGYVCKPRGFNGIHQHNH 120
QY 121 NECEAEPCRNIGICTDLVANYSCPCPGFMRNCRQYKCSGHLGIEGGIISNQOITASSNH 180
Db 121 NECEAEPCRNIGICTDLVANYSCPCPGFMRNCRQYKCSGHLGIEGGIISNQOITASSNH 180
QY 181 RALFGLQKWYPYA 194
Db 181 RALFGLQKWYPYA 194

RESULT 4
US-08-659-235C-10
; Sequence 10, Application US/08659235C
; Patent No. 5877281
; GENERAL INFORMATION:
; APPLICANT: Quettermous, Thomas
; APPLICANT: Hogan, Brigid
; APPLICANT: Snodgrass, H. Ralph
; APPLICANT: Zupancic, Thomas J.
; TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
; TITLE OF INVENTION: CELL LOCUS-1
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/659,235C
; FILING DATE: 05-JUN-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-0034-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 Pennie
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 480 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-659-235C-10

Query Match 87.8%; Score 194; DB 2; Length 480;
Best Local Similarity 100.0%; Pred. No. 5e-181;
Matches 194; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKHLVAAMLLVGLSLGVQFGKGDICNPNCENGICLSGLADDSFSCCEPGFAGPNC 60
Db 1 MKHLVAAMLLVGLSLGVQFGKGDICNPNCENGICLSGLADDSFSCCEPGFAGPNC 60

QY 61 SVVEVASDEEKPTSGPCIPNCHNGGTCEISEAYRGDTFIGYVCKPRGFNGIHQHNH 120
Db 61 SVVEVASDEEKPTSGPCIPNCHNGGTCEISEAYRGDTFIGYVCKPRGFNGIHQHNH 120

QY 121 NECEAEPCRNIGICTDLVANYSCPCPGFMRNCRQYKCSGHLGIEGGIISNQOITASSNH 180
Db 121 NECEAEPCRNIGICTDLVANYSCPCPGFMRNCRQYKCSGHLGIEGGIISNQOITASSNH 180

QY 181 RALFGLQKWYPYA 194

Db 181 RALFGLQKWYPYA 194

RESULT 5
US-08-480-229C-14
; Sequence 14, Application US/08480229C
; Patent No. 5874562
; GENERAL INFORMATION:
; APPLICANT: Quettermous, Thomas
; APPLICANT: Hogan, Brigid
; APPLICANT: Snodgrass, H. Ralph
; APPLICANT: Zupancic, Thomas J.
; TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
; TITLE OF INVENTION: CELL LOCUS-1
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,229C
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-0026-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 Pennie
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 513 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-480-229C-14

Query Match 20.4%; Score 45; DB 2; Length 513;
Best Local Similarity 100.0%; Pred. No. 5.1e-36;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 PNPCHNGTCEISEAYRGDTFIGYVCKPRGFNGIHQHNH 124
Db 112 PNPCHNGTCEISEAYRGDTFIGYVCKPRGFNGIHQHNH 156

RESULT 6
US-08-659-235C-14
; Sequence 14, Application US/08659235C
; Patent No. 5877281
; GENERAL INFORMATION:
; APPLICANT: Quettermous, Thomas
; APPLICANT: Hogan, Brigid
; APPLICANT: Snodgrass, H. Ralph
; APPLICANT: Zupancic, Thomas J.
; TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
; TITLE OF INVENTION: CELL LOCUS-1
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP

STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/659,235C
FILING DATE: 05-JUN-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-0034-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennie
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 513 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-659-235C-14

Query Match 20.4%; Score 45; DB 2; Length 513;
Best Local Similarity 100.0%; Pred. No. 5.1e-36;
Matches 45; Conservative 0; Mismatches 0; Indels 0;
Gaps 0;

QY 80 PNPCHNGTCEISEAYRGDTFIGYVCKPCPGFNGIHCOHNECE 124
Db 112 PNPCHNGTCEISEAYRGDTFIGYVCKPCPGFNGIHCOHNECE 156

RESULT 7
US-08-480-229C-24
Sequence 24, Application US/08480229C
Patent No. 5874562
GENERAL INFORMATION:
APPLICANT: Quertermous, Thomas
APPLICANT: Hogan, Brigid
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,229C
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-0026-999

TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennie
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 43 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-480-229C-24

Query Match 18.6%; Score 41; DB 2; Length 43;
Best Local Similarity 100.0%; Pred. No. 4.6e-33;
Matches 41; Conservative 0; Mismatches 0; Indels 0;
Gaps 0;

QY 80 PNPCHNGTCEISEAYRGDTFIGYVCKPCPGFNGIHCOHNI 120
Db 3 PNPCHNGTCEISEAYRGDTFIGYVCKPCPGFNGIHCOHNI 43

RESULT 8
US-08-659-235C-24
Sequence 24, Application US/08659235C
Patent No. 5877281
GENERAL INFORMATION:
APPLICANT: Quertermous, Thomas
APPLICANT: Hogan, Brigid
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: CELL LOCUS-1
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/659,235C
FILING DATE: 05-JUN-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-0034-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennie
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 43 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-659-235C-24

Query Match 18.6%; Score 41; DB 2; Length 43;
Best Local Similarity 100.0%; Pred. No. 4.6e-33;
Matches 41; Conservative 0; Mismatches 0; Indels 0;
Gaps 0;

QY 80 PNPCHNGTCEISAYRGDTFYGVCCKPRGFNGIHCOHNI 120
|||||
Db 3 PNPCHNGTCEISAYRGDTFYGVCCKPRGFNGIHCOHNI 43

RESULT 9
US-08-480-229C-22
; Sequence 22, Application US/08480229C
; Patent No. 5874562
; GENERAL INFORMATION:
; APPLICANT: Quettermous, Thomas
; APPLICANT: Hogan, Brigid
; APPLICANT: Snodgrass, H. Ralph
; APPLICANT: Zupancic, Thomas J.
; TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
; TITLE OF INVENTION: CELL LOCUS-1
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,229C
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-0026-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; MOLECULE TYPE: peptide
; US-08-480-229C-22

Query Match 8.1%; Score 18; DB 2; Length 25;
Best Local Similarity 100.0%; Pred. No. 6.8e-11;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 WLLVGLSLGVPOFGKGDI 25
|||||
Db 8 WLLVGLSLGVPOFGKGDI 25

RESULT 10
US-08-659-235C-22
; Sequence 22, Application US/08659235C
; Patent No. 5877281
; GENERAL INFORMATION:
; APPLICANT: Quettermous, Thomas
; APPLICANT: Hogan, Brigid
; APPLICANT: Snodgrass, H. Ralph
; APPLICANT: Zupancic, Thomas J.
; TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
; TITLE OF INVENTION: CELL LOCUS-1
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/659,235C
; FILING DATE: 05-JUN-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-0034-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; MOLECULE TYPE: peptide
; US-08-659-235C-22

Query Match 8.1%; Score 18; DB 2; Length 25;
Best Local Similarity 100.0%; Pred. No. 6.8e-11;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 WLLVGLSLGVPOFGKGDI 25
|||||
Db 8 WLLVGLSLGVPOFGKGDI 25

RESULT 11
US-08-480-229C-25
; Sequence 25, Application US/08480229C
; Patent No. 5874562
; GENERAL INFORMATION:
; APPLICANT: Quettermous, Thomas
; APPLICANT: Hogan, Brigid
; APPLICANT: Snodgrass, H. Ralph
; APPLICANT: Zupancic, Thomas J.
; TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
; TITLE OF INVENTION: CELL LOCUS-1
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,229C
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
; REGISTRATION NUMBER: 28,462

REFERENCE/DOCKET NUMBER: 8907-0026-999
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennie
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 42 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-480-229C-25

Query Match 8.18; Score 18; DB 2; Length 42;
Best Local Similarity 100.0%; Pred. No. 1.1e-10;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 140 NYSCECPGEGFMGRNCQYK 157
Db 25 NYSCECPGEGFMGRNCQYK 42

RESULT 12
US-08-659-235C-25
Sequence 25, Application US/08659235C
Patent No. 5877281
GENERAL INFORMATION:
APPLICANT: Quertermous, Thomas
APPLICANT: Hogan, Brigid
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
FILING DATE: 05-JUN-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-0034-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennie
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 42 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-659-235C-25

Query Match 8.18; Score 18; DB 2; Length 42;
Best Local Similarity 100.0%; Pred. No. 1.1e-10;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 140 NYSCECPGEGFMGRNCQYK 157
Db 25 NYSCECPGEGFMGRNCQYK 42

RESULT 13
US-08-480-229C-21
Sequence 21, Application US/08480229C
Patent No. 587A562
GENERAL INFORMATION:
APPLICANT: Quertermous, Thomas
APPLICANT: Hogan, Brigid
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-0026-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennie
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 321 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-480-229C-21

Query Match 8.18; Score 18; DB 2; Length 321;
Best Local Similarity 100.0%; Pred. No. 6.4e-10;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 196 LNKKGGLINAWTAENDRW 213
Db 39 LNKKGGLINAWTAENDRW 56

RESULT 14
US-08-659-235C-21
Sequence 21, Application US/08659235C
Patent No. 5877281
GENERAL INFORMATION:
APPLICANT: Quertermous, Thomas
APPLICANT: Hogan, Brigid
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
NUMBER OF SEQUENCES: 29

```
;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/659,235C
; FILING DATE: 05-JUN-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-0034-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 Pennie
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 321 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-659-235C-21

Query Match 8.1%; Score 18; DB 2; Length 321;
Best Local Similarity 100.0%; Pred. No. 6.4e-10;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 196 LNKKGGLINAWTAENDRW 213
DB 39 LNKKGGLINAWTAENDRW 56

RESULT 15
US-08-480-229C-23
; Sequence 23, Application US/08480229C
; Patent No. 5874562
; GENERAL INFORMATION:
; APPLICANT: Quertemous, Thomas
; APPLICANT: Hogan, Brigid
; APPLICANT: Snodgrass, H. Ralph
; APPLICANT: Zupancic, Thomas J.
; TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
; NUMBER OF SEQUENCES: 29
; TITLE OF INVENTION: CELL LOCUS-1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,229C
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
```

```
;
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-0026-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 Pennie
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 57 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; US-08-480-229C-23

Query Match 6.3%; Score 14; DB 2; Length 57;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 PNCSSVVEVASDEE 70
DB 37 PNCSSVVEVASDEE 50

RESULT 16
US-08-659-235C-23
; Sequence 23, Application US/08659235C
; Patent No. 5877281
; GENERAL INFORMATION:
; APPLICANT: Quertemous, Thomas
; APPLICANT: Hogan, Brigid
; APPLICANT: Snodgrass, H. Ralph
; APPLICANT: Zupancic, Thomas J.
; TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
; NUMBER OF SEQUENCES: 29
; TITLE OF INVENTION: CELL LOCUS-1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/659,235C
; FILING DATE: 05-JUN-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-0034-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 Pennie
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 57 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; US-08-659-235C-23

Query Match 6.3%; Score 14; DB 2; Length 57;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
```

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 PNCSSVVEASDEE 70

Db 37 PNCSSVVEASDEE 50

RESULT 17

US-08-882-046-7
; Sequence 7, Application US/08882046
; Patent No. 6136952
; GENERAL INFORMATION:
; APPLICANT: Li, Linheng
; APPLICANT: Hood, Leroy
; APPLICANT: Krantz, Ian D.
; APPLICANT: Spinner, Nancy B.
; TITLE OF INVENTION: Human Jagged Polypeptide, Encoding
; TITLE OF INVENTION: Nucleic Acids and Methods of Use
; NUMBER OF SEQUENCES: 110
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/882,046
; FILING DATE: 25-JUN-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-UW 2637
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1010 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-882-046-7

Query Match 4.5%; Score 10; DB 4; Length 1010;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 PNPCHNGGTC 89

Db 750 PNPCHNGGTC 759

RESULT 18

US-08-400-159-10
; Sequence 10, Application US/08400159
; Patent No. 5869282
; GENERAL INFORMATION:
; APPLICANT: Ish-Horowicz, David
; APPLICANT: Henrique, Domingos M.P.
; APPLICANT: Lewis, Julian H.
; APPLICANT: Myat, Anna M.
; APPLICANT: Fleming, Robert J.
; APPLICANT: Artavanis-Tsakonas, Spyridon
; APPLICANT: Mann, Robert S.
; APPLICANT: Gray, Grace E.

; TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF THE
; TITLE OF INVENTION: SERRATE GENE AND METHODS BASED THEREON
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/08/400,159
; FILING DATE: 07-MAR-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mirock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7326-029
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1193 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-400-159-10

Query Match 4.5%; Score 10; DB 2; Length 1193;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 PNPCHNGGTC 89

Db 724 PNPCHNGGTC 733

RESULT 19

US-08-611-729A-10
; Sequence 10, Application US/08611729A
; Patent No. 6004924
; GENERAL INFORMATION:
; APPLICANT: Ish-Horowicz, David
; APPLICANT: Henrique, Domingos M.P.
; APPLICANT: Lewis, Julian H.
; APPLICANT: Myat, Anna M.
; APPLICANT: Fleming, Robert J.
; APPLICANT: Artavanis-Tsakonas, Spyridon
; APPLICANT: Mann, Robert S.
; APPLICANT: Gray, Grace E.
; TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF THE
; TITLE OF INVENTION: SERRATE GENE AND METHODS BASED THEREON
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30

```
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/611,729A
; FILING DATE: 06-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistrock, S. Leslie
; REGISTRATION NUMBER: 16,872
; REFERENCE/DOCKET NUMBER: 7326-037
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1193 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-611-729A-10

Query Match      4.5%; Score 10; DB 3; Length 1193;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 PNPCHNGGTC 89
DB 724 PNPCHNGGTC 733

RESULT 20
US-08-400-159-6
; Sequence 6, Application US/08400159
; Patent No. 5869282
; GENERAL INFORMATION:
; APPLICANT: Ish-Horowicz, David
; APPLICANT: Henrique, Domingos M.P.
; APPLICANT: Lewis, Julian H.
; APPLICANT: Myat, Anna M.
; APPLICANT: Fleming, Robert J.
; APPLICANT: Artavanis-Tsakonas, Spyridon
; APPLICANT: Mann, Robert S.
; APPLICANT: Gray, Grace E.
; TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF THE
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/400,159
; FILING DATE: 07-MAR-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistrock, S. Leslie
; REGISTRATION NUMBER: 16,872
; REFERENCE/DOCKET NUMBER: 7326-029
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1218 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-611-729A-6

Query Match      4.5%; Score 10; DB 3; Length 1218;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 PNPCHNGGTC 89
DB 750 PNPCHNGGTC 759
```

```
;
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-400-159-6

Query Match      4.5%; Score 10; DB 2; Length 1218;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 PNPCHNGGTC 89
DB 750 PNPCHNGGTC 759

RESULT 21
US-08-611-729A-6
; Sequence 6, Application US/08611729A
; Patent No. 6004924
; GENERAL INFORMATION:
; APPLICANT: Ish-Horowicz, David
; APPLICANT: Henrique, Domingos M.P.
; APPLICANT: Lewis, Julian H.
; APPLICANT: Myat, Anna M.
; APPLICANT: Fleming, Robert J.
; APPLICANT: Artavanis-Tsakonas, Spyridon
; APPLICANT: Mann, Robert S.
; APPLICANT: Gray, Grace E.
; TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF THE
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/611,729A
; FILING DATE: 06-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistrock, S. Leslie
; REGISTRATION NUMBER: 16,872
; REFERENCE/DOCKET NUMBER: 7326-037
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1218 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-611-729A-6

Query Match      4.5%; Score 10; DB 3; Length 1218;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 PNPCHNGGTC 89
DB 750 PNPCHNGGTC 759
```

RESULT 22
US-08-882-046-2
; Sequence 2, Application US/08882046
; Patent No. 6136952
; GENERAL INFORMATION:
; APPLICANT: Li, Linheng
; APPLICANT: Hood, Leroy
; APPLICANT: Krantz, Ian D.
; APPLICANT: Spinner, Nancy B.
; TITLE OF INVENTION: Human Jagged Polypeptide, Encoding
; TITLE OF INVENTION: Nucleic Acids and Methods of Use
; NUMBER OF SEQUENCES: 110
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/882,046
; FILING DATE: 25-JUN-1997
; CLASSIFICATION: 536
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-UW 2637
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1218 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-882-046-2

Query Match 4.5%; Score 10; DB 4; Length 1218;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 PNPCHNGGTC 89
|||||
DB 750 PNPCHNGGTC 759

RESULT 23
US-08-882-046-5
; Sequence 5, Application US/08882046
; Patent No. 6136952
; GENERAL INFORMATION:
; APPLICANT: Li, Linheng
; APPLICANT: Hood, Leroy
; APPLICANT: Krantz, Ian D.
; APPLICANT: Spinner, Nancy B.
; TITLE OF INVENTION: Human Jagged Polypeptide, Encoding
; TITLE OF INVENTION: Nucleic Acids and Methods of Use
; NUMBER OF SEQUENCES: 110
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/882,046
; FILING DATE: 25-JUN-1997
; CLASSIFICATION: 536
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-UW 2637
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1219 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-882-046-5

Query Match 4.5%; Score 10; DB 4; Length 1219;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 PNPCHNGGTC 89
|||||
DB 750 PNPCHNGGTC 759

RESULT 24
US-08-185-432-17
; Sequence 17, Application US/08185432
; Patent No. 5750852
; GENERAL INFORMATION:
; APPLICANT: Artavanis-Tsakonas, Spyridon
; APPLICANT: Busseau, Isabelle
; APPLICANT: Diederich, Robert J.
; APPLICANT: Xu, Tian
; APPLICANT: Matsuno, Kenji
; TITLE OF INVENTION: DELTEX PROTEINS, NUCLEIC ACIDS, AND
; TITLE OF INVENTION: ANTIBODIES, AND RELATED METHODS AND COMPOSITIONS
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/185,432
; FILING DATE: 21-JAN-1994
; CLASSIFICATION: 530
; NAME: Mistrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7326-006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2556 amino acids

; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-185-432-17

Query Match 4.5%; Score 10; DB 1; Length 2556;
Best Local Similarity 100.0%; Pred. No. 0.24;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 81 NPCHNGGTCE 90
Db 685 NPCHNGGTCE 694

RESULT 25
US-08-083-590A-20
; Sequence 20, Application US/08083590A
; Patent No. 5786158
; GENERAL INFORMATION:
; APPLICANT: Artavanis-Tsakonas, S. et al.
; TITLE OF INVENTION: Therapeutic And Diagnostic Methods
; TITLE OF INVENTION: And Compositions Based On No. 5786158ch Proteins And
; TITLE OF INVENTION: Nucleic Acids
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/083,590A
; FILING DATE: 25-JUN-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mirock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7326-015
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 790-9090
; TELEFAX: 212 8698864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2556 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-083-590A-20

Query Match 4.5%; Score 10; DB 1; Length 2556;
Best Local Similarity 100.0%; Pred. No. 0.24;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 81 NPCHNGGTCE 90
Db 685 NPCHNGGTCE 694

Search completed: May 23, 2001, 06:21:44
Job time: 383 sec

